CORRECTED VERSION*

# PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

AL

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)


(51) International Patent Classification 7 : C12N 15/12, C07K 14/72, G01N 33/50, 33/566 — A1

(11) International Publication Number: WO 00/22129

(43) International Publication Date: 20 April 2000 (20.04.00)

(21) International Application Number: PCT/US99/23938

(22) International Filing Date: 12 October 1999 (12.10.99)

(30) Priority Data:
09/170,496 13 October 1998 (13.10.98) US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application
US 09/170,496 (CIP)
Filed on 13 October 1998 (13.10.98)



(71) Applicant *(for all designated States except US)*: ARENA PHARMACEUTICALS, INC. [US/US]; 6166 Nancy Ridge Drive, San Diego, CA 92121 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: BEHAN, Dominic, P. [GB/US]; 11472 Roxboro Court, San Diego, CA 92131 (US). CHALMERS, Derek, T. [GB/US]; 347 Longden Lane, Solana Beach, CA 92075 (US). LIAW, Chen, W. [US/US]; 7668 Salix Place, San Diego, CA 92129 (US).

(74) Agents: MILLER, Suzanne, E. et al.; Woodcock Washburn Kurtz Mackiewicz & Norris LLP, 46th floor, One Liberty Place, Philadelphia, PA 19103 (US).



(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published**
*With international search report.*
*Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.*


(54) Title: NON-ENDOGENOUS, CONSTITUTIVELY ACTIVATED HUMAN G PROTEIN-COUPLED RECEPTORS


(57) Abstract

Disclosed herein are constitutively activated, non-endogenous versions of endogenous human G protein-coupled receptors comprising (a) the following amino acid sequence region (C-terminus to N-terminus orientation) and/or (b) the following nucleic acid sequence region (3' to 5' orientation) transversing the transmembrane-6 (TM6) and intracellular loop-3 (IC3) regions of the GPCR: (a) $P^1$ $AA_{15}$ X and/or (b) $P^{codon}$ $(AA\text{-}codon)_{15}$ $X_{codon}$, respectively. In a most preferred embodiment, $P^1$ and $P^{codon}$ are endogenous proline and an endogenous nucleic acid encoding region encoding proline, respectively; located within TM6 of the non-endogenous GPCR; $AA_{15}$ and $(AA\text{-}codon)_{15}$ are 15 endogenous amino acid residues and 15 codons encoding endogenous amino acid residues, respectively; and X and $X_{codon}$ are non-endogenous lysine and a non-endogenous nucleic acid encoding region encoding lysine, respectively, located within IC3 of the non-endogenous GPCR. Because it is most preferred that the non-endogenous human GPCRs which incorporate these mutations are incorporated into mammalian cells and utilized for the screening of the candidate compounds, the non-endogenous human GPCR incorporating the mutation need not be purified and isolated *per se* (i.e., these are incorporated within the cellular membrane of a mammalian cell), although such purified and isolated non-endogenous human GPCRs are well within the purview of this disclosure.


*(Referred to in PCT Gazette No. 37/2000, Section II)

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| AL | Albania | ES | Spain | LS | Lesotho | SI | Slovenia |
| AM | Armenia | FI | Finland | LT | Lithuania | SK | Slovakia |
| AT | Austria | FR | France | LU | Luxembourg | SN | Senegal |
| AU | Australia | GA | Gabon | LV | Latvia | SZ | Swaziland |
| AZ | Azerbaijan | GB | United Kingdom | MC | Monaco | TD | Chad |
| BA | Bosnia and Herzegovina | GE | Georgia | MD | Republic of Moldova | TG | Togo |
| BB | Barbados | GH | Ghana | MG | Madagascar | TJ | Tajikistan |
| BE | Belgium | GN | Guinea | MK | The former Yugoslav Republic of Macedonia | TM | Turkmenistan |
| BF | Burkina Faso | GR | Greece | | | TR | Turkey |
| BG | Bulgaria | HU | Hungary | ML | Mali | TT | Trinidad and Tobago |
| BJ | Benin | IE | Ireland | MN | Mongolia | UA | Ukraine |
| BR | Brazil | IL | Israel | MR | Mauritania | UG | Uganda |
| BY | Belarus | IS | Iceland | MW | Malawi | US | United States of America |
| CA | Canada | IT | Italy | MX | Mexico | UZ | Uzbekistan |
| CF | Central African Republic | JP | Japan | NE | Niger | VN | Viet Nam |
| CG | Congo | KE | Kenya | NL | Netherlands | YU | Yugoslavia |
| CH | Switzerland | KG | Kyrgyzstan | NO | Norway | ZW | Zimbabwe |
| CI | Côte d'Ivoire | KP | Democratic People's Republic of Korea | NZ | New Zealand | | |
| CM | Cameroon | | | PL | Poland | | |
| CN | China | KR | Republic of Korea | PT | Portugal | | |
| CU | Cuba | KZ | Kazakstan | RO | Romania | | |
| CZ | Czech Republic | LC | Saint Lucia | RU | Russian Federation | | |
| DE | Germany | LI | Liechtenstein | SD | Sudan | | |
| DK | Denmark | LK | Sri Lanka | SE | Sweden | | |
| EE | Estonia | LR | Liberia | SG | Singapore | | |

# NON-ENDOGENOUS, CONSTITUTIVELY ACTIVATED HUMAN G PROTEIN-COUPLED RECEPTORS

The benefits of commonly owned U.S. Serial Number 09/170,496, filed October 13, 1998, U.S. Serial Number 08/839, 449 filed April 14, 1997 (now abandoned), U.S. Serial Number 09/060,188, filed April 14, 1998; U.S. Provisional Number 60/090,783, filed June 26, 1998; and U.S. Provisional Number 60/095,677, filed on August 7, 1998, are hereby claimed. Each of the foregoing applications are incorporated by reference herein in their entirety.

## FIELD OF THE INVENTION

The invention disclosed in this patent document relates to transmembrane receptors, and more particularly to human G protein-coupled receptors (GPCRs) which have been altered such that altered GPCRs are constitutively activated. Most preferably, the altered human GPCRs are used for the screening of therapeutic compounds.

## BACKGROUND OF THE INVENTION

Although a number of receptor classes exist in humans, by far the most abundant and therapeutically relevant is represented by the G protein-coupled receptor (GPCR or GPCRs) class. It is estimated that there are some 100,000 genes within the human genome, and of these, approximately 2% or 2,000 genes, are estimated to code for GPCRs. Of these, there are approximately 100 GPCRs for which the endogenous ligand that binds to the GPCR has been identified. Because of the significant time-lag that exists between the discovery of an endogenous GPCR and its endogenous ligand, it can be presumed that the remaining 1,900 GPCRs will be identified and characterized long before the endogenous ligands for these receptors are identified. Indeed, the rapidity by which the Human Genome Project is sequencing the 100,000 human genes indicates that the remaining human GPCRs will be fully sequenced within the next few years. Nevertheless, and despite the efforts to sequence the human genome, it is still very unclear as to how scientists will be able to rapidly, effectively and efficiently exploit this information to improve and enhance the human condition. The present invention is geared towards this important objective.

Receptors, including GPCRs, for which the endogenous ligand has been identified are referred to as "known" receptors, while receptors for which the endogenous ligand has not been identified are referred to as "orphan" receptors. This distinction is not merely semantic, particularly in the case of GPCRs. GPCRs represent an important area for the development of pharmaceutical products: from approximately 20 of the 100 known GPCRs, 60% of all prescription pharmaceuticals have been developed. Thus, the orphan GPCRs are to the pharmaceutical industry what gold was to California in the late 19$^{th}$ century – an opportunity to drive growth, expansion, enhancement and development. A serious drawback exists, however,

with orphan receptors relative to the discovery of novel therapeutics. This is because the traditional approach to the discovery and development of pharmaceuticals has required access to both the receptor *and* its endogenous ligand. Thus, heretofore, orphan GPCRs have presented the art with a tantalizing and undeveloped resource for the discovery of pharmaceuticals.

Under the traditional approach to the discovery of potential therapeutics, it is generally the case that the receptor is first identified. Before drug discovery efforts can be initiated, elaborate, time consuming and expensive procedures are typically put into place in order to identify, isolate and generate the receptor's endogenous ligand – this process can require from between 3 and ten years per receptor, at a cost of about $5million (U.S.) per receptor. These time and financial resources must be expended before the traditional approach to drug discovery can commence. This is because traditional drug discovery techniques rely upon so-called "competitive binding assays" whereby putative therapeutic agents are "screened" against the receptor in an effort to discover compounds that either block the endogenous ligand from binding to the receptor ("antagonists"), or enhance or mimic the effects of the ligand binding to the receptor ("agonists"). The overall objective is to identify compounds that prevent cellular activation when the ligand binds to the receptor (the antagonists), or that enhance or increase cellular activity that would otherwise occur if the ligand was properly binding with the receptor (the agonists). Because the endogenous ligands for orphan GPCRs are by definition not identified, the ability to discover novel and unique therapeutics to these receptors using traditional drug discovery techniques is not possible. The present invention, as will be set forth in greater detail below, overcomes these and other severe limitations created by such traditional drug discovery techniques.

GPCRs share a common structural motif. All these receptors have seven sequences of between 22 to 24 hydrophobic amino acids that form seven alpha helices, each of which spans the

membrane (each span is identified by number, *i.e.*, transmembrane-1 (TM-1), transmebrane-2 (TM-2), etc.). The transmembrane helices are joined by strands of amino acids between transmembrane-2 and transmembrane-3, transmembrane-4 and transmembrane-5, and transmembrane-6 and transmembrane-7 on the exterior, or "extracellular" side, of the cell membrane (these are referred to as "extracellular" regions 1, 2 and 3 (EC-1, EC-2 and EC-3), respectively). The transmembrane helices are also joined by strands of amino acids between transmembrane-1 and transmembrane-2, transmembrane-3 and transmembrane-4, and transmembrane-5 and transmembrane-6 on the interior, or "intracellular" side, of the cell membrane (these are referred to as "intracellular" regions 1, 2 and 3 (IC-1, IC-2 and IC-3), respectively). The "carboxy" ("C") terminus of the receptor lies in the intracellular space within the cell, and the "amino" ("N") terminus of the receptor lies in the extracellular space outside of the cell. The general structure of G protein-coupled receptors is depicted in Figure 1.

Generally, when an endogenous ligand binds with the receptor (often referred to as "activation" of the receptor), there is a change in the conformation of the intracellular region that allows for coupling between the intracellular region and an intracellular "G-protein." Although other G proteins exist, currently, Gq, Gs, Gi, and Go are G proteins that have been identified. Endogenous ligand-activated GPCR coupling with the G-protein begins a signaling cascade process (referred to as "signal transduction"). Under normal conditions, signal transduction ultimately results in cellular activation or cellular inhibition. It is thought that the IC-3 loop as well as the carboxy terminus of the receptor interact with the G protein. A principal focus of this invention is directed to the transmembrane-6 (TM6) region and the intracellular-3 (IC3) region of the GPCR.

Under physiological conditions, GPCRs exist in the cell membrane in equilibrium between

two different conformations: an "inactive" state and an "active" state. As shown schematically in Figure 2, a receptor in an inactive state is unable to link to the intracellular signaling transduction pathway to produce a biological response. Changing the receptor conformation to the active state allows linkage to the transduction pathway (via the G-protein) and produces a biological response.

A receptor may be stabilized in an active state by an endogenous ligand or a compound such as a drug. Recent discoveries, including but not exclusively limited to modifications to the amino acid sequence of the receptor, provide means other than endogenous ligands or drugs to promote and stabilize the receptor in the active state conformation. These means effectively stabilize the receptor in an active state by simulating the effect of an endogenous ligand binding to the receptor. Stabilization by such ligand-independent means is termed "constitutive receptor activation."

As noted above, the use of an orphan receptor for screening purposes has not been possible. This is because the traditional "dogma" regarding screening of compounds mandates that the ligand for the receptor be known. By definition, then, this approach has no applicability with respect to orphan receptors. Thus, by adhering to this dogmatic approach to the discovery of therapeutics, the art, in essence, has taught and has been taught to forsake the use of orphan receptors unless and until the endogenous ligand for the receptor is discovered. Given that there are an estimated 2,000 G protein coupled receptors, the majority of which are orphan receptors, such dogma castigates a creative, unique and distinct approach to the discovery of therapeutics.

Information regarding the nucleic acid and/or amino acid sequences of a variety of GPCRs is summarized below in Table A. Because an important focus of the invention disclosed herein is directed towards orphan GPCRs, many of the below-cited references are related to orphan GPCRs. However, this list is not intended to imply, nor is this list to be construed, legally or

otherwise, that the invention disclosed herein is only applicable to orphan GPCRs or the specific GPCRs listed below. Additionally, certain receptors that have been isolated are not the subject of publications per se; for example, reference is made to a G Protein-Coupled Receptor database on the "world-wide web" (neither the named inventors nor the assignee have any affiliation with this site) that lists GPCRs. Other GPCRs are the subject of patent applications owned by the present assignee and these are not listed below (including GPR3, GPR6 and GPR12; *see* U.S. Provisional Number 60/094879):

**Table A**

| Receptor Name | Publication Reference |
|---|---|
| GPR1 | 23 Genomics 609 (1994) |
| GPR4 | 14 DNA and Cell Biology 25 (1995) |
| GPR5 | 14 DNA and Cell Biology 25 (1995) |
| GPR7 | 28 Genomics 84 (1995) |
| GPR8 | 28 Genomics 84 (1995) |
| GPR9 | 184 J. Exp. Med. 963 (1996) |
| GPR10 | 29 Genomics 335 (1995) |
| GPR15 | 32 Genomics 462 (1996) |
| GPR17 | 70 J Neurochem. 1357 (1998) |
| GPR18 | 42 Genomics 462 (1997) |
| GPR20 | 187 Gene 75 (1997) |
| GPR21 | 187 Gene 75 (1997) |
| GPR22 | 187 Gene 75 (1997) |
| GPR24 | 398 FEBS Lett. 253 (1996) |
| GPR30 | 45 Genomics 607 (1997) |
| GPR31 | 42 Genomics 519 (1997) |
| GPR32 | 50 Genomics 281 (1997) |
| GPR40 | 239 Biochem. Biophys. Res. Commun. 543 (1997) |
| GPR41 | 239 Biochem. Biophys. Res. Commun. 543 (1997) |
| GPR43 | 239 Biochem. Biophys. Res. Commun. 543 (1997) |
| APJ | 136 Gene 355 (1993) |
| BLR1 | 22 Eur. J. Immunol. 2759 (1992) |
| CEPR | 231 Biochem. Biophys. Res. Commun. 651 (1997) |
| EBI1 | 23 Genomics 643 (1994) |
| EBI2 | 67 J. Virol. 2209 (1993) |
| ETBR-LP2 | 424 FEBS Lett. 193 (1998) |
| GPCR-CNS | 54 Brain Res. Mol. Brain Res. 152 (1998); 45 Genomics 68 (1997) |
| GPR-NGA | 394 FEBS Lett. 325 (1996) |
| H9 | 386 FEBS Lett 219 (1996) |

| | |
|---|---|
| HBA954 | 1261 Biochim. Biophys. Acta 121 (1995) |

| | |
|---|---|
| HG38 | 247 Biochem. Biophys. Res. Commun. 266 (1998) |
| HM74 | 5 Int. Immunol. 1239 (1993) |
| OGR1 | 35 Genomics 397 (1996) |
| V28 | 163 Gene 295 (1995) |

As will be set forth and disclosed in greater detail below, utilization of a mutational cassette to modify the endogenous sequence of a human GPCR leads to a constitutively activated version of the human GPCR. These non-endogenous, constitutively activated versions of human GPCRs can be utilized, *inter alia*, for the screening of candidate compounds to directly identify compounds of, *e.g.*, therapeutic relevance.

## SUMMARY OF THE INVENTION

Disclosed herein is a non-endogenous, human G protein-coupled receptor comprising (a) as a most preferred amino acid sequence region (C-terminus to N-terminus orientation) and/or (b) as a most preferred nucleic acid sequence region (3' to 5' orientation) transversing the transmembrane-6 (TM6) and intracellular loop-3 (IC3) regions of the GPCR:

(a) $P^1\ AA_{15}\ X$

wherein:

(1) $P^1$ is an amino acid residue located within the TM6 region of the GPCR, where $P^1$ is selected from the group consisting of (i) the endogenous GPCR's proline residue, and (ii) a non-endogenous amino acid residue other than proline;

(2) $AA_{15}$ are 15 amino acids selected from the group consisting of

(a) the endogenous GPCR's amino acids (b) non-endogenous amino acid residues, and (c) a combination of the endogenous GPCR's amino acids and non-endogenous amino acids, excepting that none of the 15 endogenous amino acid residues that are positioned within the TM6 region of the GPCR is proline; and

(3) X is a non-endogenous amino acid residue located within the IC3 region of said GPCR, preferably selected from the group consisting of lysine, hisitidine and arginine, and most preferably lysine, excepting that when the endogenous amino acid at position X is lysine, then X is an amino acid other than lysine, preferably alanine;

and/or

(b) $P^{codon}$ $(AA\text{-}codon)_{15}$ $X_{codon}$

wherein:

(1) $P^{codon}$ is a nucleic acid sequence within the TM6 region of the GPCR, where $P^{codon}$ encodes an amino acid selected from the group consisting of (i) the endogenous GPCR's proline residue, and (ii) a non-endogenous amino acid residue other than proline;

(2) $(AA\text{-}codon)_{15}$ are 15 codons encoding 15 amino acids selected from the group consisting of (a) the endogenous GPCR's amino acids (b) non-endogenous amino acid residues and (c) a combination of the endogenous GPCR's amino acids and non-

endogenous amino acids, excepting that none of the 15 endogenous codons within the TM6 region of the GPCR encodes a proline amino acid residue; and

(3) $X_{codon}$ is a nucleic acid encoding region residue located within the IC3 region of said GPCR, where $X_{codon}$ encodes a non-endogenous amino acid, preferably selected from the group consisting of lysine, hisitidine and arginine, and most preferably lysine, excepting that when the endogenous encoding region at position $X_{codon}$ encodes the amino acid lysine, then $X_{codon}$ encodes an amino acid other than lysine, preferably alanine.

The terms endogenous and non-endogenous in reference to these sequence cassettes are relative to the endogenous GPCR. For example, once the endogenous proline residue is located within the TM6 region of a particular GPCR, and the 16th amino acid therefrom is identified for mutation to constitutively activate the receptor, it is also possible to mutate the endogenous proline residue (*i.e.*, once the marker is located and the 16th amino acid to be mutated is identified, one may mutate the marker itself), although it is most preferred that the proline residue not be mutated. Similarly, and while it is most preferred that $AA_{15}$ be maintained in their endogenous forms, these amino acids may also be mutated. The only amino acid that must be mutated in the non-endogenous version of the human GPCR is X *i.e.*, the endogenous amino acid that is 16 residues from $P^1$ cannot be maintained in its endogenous form and must be mutated, as further disclosed herein. Stated again, while it is preferred that in the non-endogenous version of the human GPCR, $P^1$ and $AA_{15}$ remain in their endogenous forms (*i.e.*, identical to their wild-type forms), once X is identified and mutated, any and/or all of $P^1$ and $AA_{15}$ can be mutated. This applies to the nucleic

acid sequences as well. In those cases where the endogenous amino acid at position X is lysine, then in the non-endogenous version of such GPCR, X is an amino acid other than lysine, preferably alanine.

Accordingly, and as a hypothetical example, if the endogenous GPCR has the following endogenous amino acid sequence at the above-noted positions:

P-AACCTTGGRRRDDDE -Q

then any of the following exemplary and hypothetical cassettes would fall within the scope of the disclosure (non-endogenous amino acids are set forth in bold):

P-AACCTTGGRRRDDDE -**K**

P-AACCTT**HI**GRRDDDE -**K**

P-**ADEE**TTGGRRRDDDE -**A**

**P-LLKFMSTWZLVAAPQ -K**

**A-LLKFMSTWZLVAAPQ -K**

It is also possible to add amino acid residues within $AA_{15}$, but such an approach is not particularly advanced. Indeed, in the most preferred embodiments, the only amino acid that differs in the non-endogenous version of the human GPCR as compared with the endogenous version of that GPCR is the amino acid in position X; mutation of this amino acid itself leads to constitutive activation of the receptor.

Thus, in particularly preferred embodiments, $P^1$ and $P^{codon}$ are endogenous proline and an endogenous nucleic acid encoding region encoding proline, respectively; and X and $X_{codon}$ are non-endogenous lysine or alanine and a non-endogenous nucleic acid encoding region encoding lysine or alanine, respectively, with lysine being most preferred. Because it is most preferred that the non-endogenous versions of the human GPCRs which incorporate these mutations are

incorporated into mammalian cells and utilized for the screening of candidate compounds, the non-endogenous human GPCR incorporating the mutation need not be purified and isolated *per se* (*i.e.*, these are incorporated within the cellular membrane of a mammalian cell), although such purified and isolated non-endogenous human GPCRs are well within the purview of this disclosure. Gene-targeted and transgenic non-human mammals (preferably rats and mice) incorporating the non-endogenous human GPCRs are also within the purview of this invention; in particular, gene-targeted mammals are most preferred in that these animals will incorporate the non-endogenous versions of the human GPCRs in place of the non-human mammal's endogenous GPCR-encoding region (techniques for generating such non-human mammals to replace the non-human mammal's protein encoding region with a human encoding region are well known; see, for example, U.S. Patent No. 5,777,194.)

It has been discovered that these changes to an endogenous human GPCR render the GPCR constitutively active such that, as will be further disclosed herein, the non-endogenous, constitutively activated version of the human GPCR can be utilized for, *inter alia*, the direct screening of candidate compounds without the need for the endogenous ligand. Thus, methods for using these materials, and products identified by these methods are also within the purview of the following disclosure.

## BRIEF DESCRIPTION OF THE DRAWINGS

**Figure 1** shows a generalized structure of a G protein-coupled receptor with the numbers assigned to the transmembrane helixes, the intracellular loops, and the extracellular loops.

**Figure 2** schematically shows the two states, active and inactive, for a typical G protein coupled receptor and the linkage of the active state to the second messenger transduction pathway.

**Figure 3** is a sequence diagram of the preferred vector pCMV, including restriction enzymen site locations.

**Figure 4** is a diagrammatic representation of the signal measured comparing pCMV, non-endogenous, constitutively active GPR30 inhibition of GPR6-mediated activation of CRE-Luc reporter with endogenous GPR30 inhibition of GPR6-mediated activation of CRE-Luc reporter.

**Figure 5** is a diagrammatic representation of the signal measured comparing pCMV, non-endogenous, constitutively activated GPR17 inhibition of GPR3-mediated activation of CRE-Luc reporter with endogenous GPR17 inhibition of GPR3-mediated activation of CRE-Luc reporter.

**Figure 6** provides diagrammatic results of the signal measured comparing control pCMV, endogenous APJ and non-endogenous APJ.

**Figure 7** provides an illustration of $IP_3$ production from non-endogenous human 5-$HT_{2A}$ receptor as compared to the endogenous version of this receptor.

**Figure 8** are dot-blot format results for GPR1 (8A), GPR30 (8B) and APJ (8C).

## DETAILED DESCRIPTION

The scientific literature that has evolved around receptors has adopted a number of terms to refer to ligands having various effects on receptors. For clarity and consistency, the following definitions will be used throughout this patent document. To the extent that these definitions conflict with other definitions for these terms, the following definitions shall control:

**AGONISTS** shall mean compounds that activate the intracellular response when they bind to the receptor, or enhance GTP binding to membranes.

**AMINO ACID ABBREVIATIONS** used herein are set below:

| | | |
|---|---|---|
| ALANINE | ALA | A |
| ARGININE | ARG | R |
| ASPARAGINE | ASN | N |
| ASPARTIC ACID | ASP | D |
| CYSTEINE | CYS | C |
| GLUTAMIC ACID | GLU | E |
| GLUTAMINE | GLN | Q |
| GLYCINE | GLY | G |
| HISTIDINE | HIS | H |
| ISOLEUCINE | ILE | I |
| LEUCINE | LEU | L |
| LYSINE | LYS | K |
| METHIONINE | MET | M |
| PHENYLALANINE | PHE | F |
| PROLINE | PRO | P |
| SERINE | SER | S |
| THREONINE | THR | T |
| TRYPTOPHAN | TRP | W |
| TYROSINE | TYR | Y |
| VALINE | VAL | V |

**PARTIAL AGONISTS** shall mean compounds which activate the intracellular response when they bind to the receptor to a lesser degree/extent than do agonists, or enhance GTP binding to membranes to a lesser degree/extent than do agonists

**ANTAGONIST** shall mean compounds that competitively bind to the receptor at the same site as the agonists but which do not activate the intracellular response initiated by the active form of the receptor, and can thereby inhibit the intracellular responses by agonists or partial agonists. ANTAGONISTS do not diminish the baseline intracellular response in the absence of an agonist or partial agonist.

**CANDIDATE COMPOUND** shall mean a molecule (for example, and not limitation, a chemical compound) which is amenable to a screening technique. Preferably, the phrase

"candidate compound" does not include compounds which were publicly known to be compounds selected from the group consisting of inverse agonist, agonist or antagonist to a receptor, as previously determined by an indirect identification process ("indirectly identified compound"); more preferably, not including an indirectly identified compound which has previously been determined to have therapeutic efficacy in at least one mammal; and, most preferably, not including an indirectly identified compound which has previously been determined to have therapeutic utility in humans.

**CODON** shall mean a grouping of three nucleotides (or equivalents to nucleotides) which generally comprise a nucleoside (adenosine (A), guanosine (G), cytidine (C), uridine (U) and thymidine (T)) coupled to a phosphate group and which, when translated, encodes an amino acid.

**COMPOUND EFFICACY** shall mean a measurement of the ability of a compound to inhibit or stimulate receptor functionality, as opposed to receptor binding affinity. A preferred means of detecting compound efficacy is via measurement of, *e.g.*, [$^{35}$S]GTPγS binding, as further disclosed in the Example section of this patent document.

**CONSTITUTIVELY ACTIVATED RECEPTOR** shall mean a receptor subject to constitutive receptor activation. In accordance with the invention disclosed herein, a non-endogenous, human constitutively activated G protein-coupled receptor is one that has been mutated to include the amino acid cassette $P^{1}AA_{15}X$, as set forth in greater detail below.

**CONSTITUTIVE RECEPTOR ACTIVATION** shall mean stabilization of a receptor in the active state by means other than binding of the receptor with its endogenous ligand or a chemical equivalent thereof. Preferably, a G protein-coupled receptor subjected to constitutive receptor activation in accordance with the invention disclosed herein evidences at least a 10% difference in response (increase or decrease, as the case may be) to the signal measured for

constitutive activation as compared with the endogenous form of that GPCR, more preferably, about a 25% difference in such comparative response, and most preferably about a 50% difference in such comparative response. When used for the purposes of directly identifying candidate compounds, it is most preferred that the signal difference be at least about 50% such that there is a sufficient difference between the endogenous signal and the non-endogenous signal to differentiate between selected candidate compounds. In most instances, the "difference" will be an increase in signal; however, with respect to Gs-coupled GPCRS, the "difference" measured is preferably a decrease, as will be set forth in greater detail below.

**CONTACT** or **CONTACTING** shall mean bringing at least two moieties together, whether in an in vitro system or an in vivo system.

**DIRECTLY IDENTIFYING** or **DIRECTLY IDENTIFIED**, in relationship to the phrase "candidate compound", shall mean the screening of a candidate compound against a constitutively activated G protein-coupled receptor, and assessing the compound efficacy of such compound. This phrase is, under no circumstances, to be interpreted or understood to be encompassed by or to encompass the phrase "indirectly identifying" or "indirectly identified."

**ENDOGENOUS** shall mean a material that is naturally produced by the genome of the species. ENDOGENOUS in reference to, for example and not limitation, GPCR, shall mean that which is naturally produced by a human, an insect, a plant, a bacterium, or a virus. By contrast, the term **NON-ENDOGENOUS** in this context shall mean that which is not naturally produced by the genome of a species. For example, and not limitation, a receptor which is not constitutively active in its endogenous form, but when mutated by using the cassettes disclosed herein and thereafter becomes constitutively active, is most preferably referred to herein as a "non-endogenous, constitutively activated receptor." Both terms can be utilized to describe both "in

vivo" and "in vitro" systems. For example, and not limitation, in a screening approach, the endogenous or non-endogenous receptor may be in reference to an in vitro screening system whereby the receptor is expressed on the cell-surface of a mammalian cell. As a further example and not limitation, where the genome of a mammal has been manipulated to include a non-endogenous constitutively activated receptor, screening of a candidate compound by means of an in vivo system is viable.

**HOST CELL** shall mean a cell capable of having a Plasmid and/or Vector incorporated therein. In the case of a prokaryotic Host Cell, a Plasmid is typically replicated as an autonomous molecule as the Host Cell replicates (generally, the Plasmid is thereafter isolated for introduction into a eukaryotic Host Cell); in the case of a eukaryotic Host Cell, a Plasmid is integrated into the cellular DNA of the Host Cell such that when the eukaryotic Host Cell replicates, the Plasmid replicates. Preferably, for the purposes of the invention disclosed herein, the Host Cell is eukaryotic, more preferably, mammalian, and most preferably selected from the group consisting of 293, 293T and COS-7 cells.

**INDIRECTLY IDENTIFYING** or **INDIRECTLY IDENTIFIED** means the traditional approach to the drug discovery process involving identification of an endogenous ligand specific for an endogenous receptor, screening of candidate compounds against the receptor for determination of those which interfere and/or compete with the ligand-receptor interaction, and assessing the efficacy of the compound for affecting at least one second messenger pathway associated with the activated receptor.

**INHIBIT** or **INHIBITING**, in relationship to the term "response" shall mean that a response is decreased or prevented in the presence of a compound as opposed to in the absence of the compound.

**INVERSE AGONISTS** shall mean compounds which bind to either the endogenous form of the receptor or to the constitutively activated form of the receptor, and which inhibit the baseline intracellular response initiated by the active form of the receptor below the normal base level of activity which is observed in the absence of agonists or partial agonists, or decrease GTP binding to membranes. Preferably, the baseline intracellular response is inhibited in the presence of the inverse agonist by at least 30%, more preferably by at least 50%, and most preferably by at least 75%, as compared with the baseline response in the absence of the inverse agonist.

**KNOWN RECEPTOR** shall mean an endogenous receptor for which the endogenous ligand specific for that receptor has been identified.

**LIGAND** shall mean an endogenous, naturally occurring molecule specific for an endogenous, naturally occurring receptor.

**MUTANT** or **MUTATION** in reference to an endogenous receptor's nucleic acid and/or amino acid sequence shall mean a specified change or changes to such endogenous sequences such that a mutated form of an endogenous, non-constitutively activated receptor evidences constitutive activation of the receptor. In terms of equivalents to specific sequences, a subsequent mutated form of a human receptor is considered to be equivalent to a first mutation of the human receptor if (a) the level of constitutive activation of the subsequent mutated form of the receptor is substantially the same as that evidenced by the first mutation of the receptor; and (b) the percent sequence (amino acid and/or nucleic acid) homology between the subsequent mutated form of the receptor and the first mutation of the receptor is at least about 80%, more preferably at least about 90% and most preferably at least 95%. Ideally, and owing to the fact that the most preferred cassettes disclosed herein for achieving constitutive activation includes a single amino acid and/or codon change between the endogenous and the non-endogenous forms of the GPCR (i.e. X or

$X_{codon}$), the percent sequence homology should be at least 98%.

**ORPHAN RECEPTOR** shall mean an endogenous receptor for which the endogenous ligand specific for that receptor has not been identified or is not known.

**PHARMACEUTICAL COMPOSITION** shall mean a composition comprising at least one active ingredient, whereby the composition is amenable to investigation for a specified, efficacious outcome in a mammal (for example, and not limitation, a human). Those of ordinary skill in the art will understand and appreciate the techniques appropriate for determining whether an active ingredient has a desired efficacious outcome based upon the needs of the artisan.

**PLASMID** shall mean the combination of a Vector and cDNA. Generally, a Plasmid is introduced into a Host Cell for the purpose of replication and/or expression of the cDNA as a protein.

**STIMULATE** or **STIMULATING**, in relationship to the term "response" shall mean that a response is increased in the presence of a compound as opposed to in the absence of the compound.

**TRANSVERSE** or **TRANSVERSING**, in reference to either a defined nucleic acid sequence or a defined amino acid sequence, shall mean that the sequence is located within at least two different and defined regions. For example, in an amino acid sequence that is 10 amino acid moieties in length, where 3 of the 10 moieties are in the TM6 region of a GPCR and the remaining 7 moieties are in the IC3 region of the GPCR, the 10 amino acid moiety can be described as transversing the TM6 and IC3 regions of the GPCR.

**VECTOR** in reference to cDNA shall mean a circular DNA capable of incorporating at least one cDNA and capable of incorporation into a Host Cell.

The order of the following sections is set forth for presentational efficiency and is not

intended, nor should be construed, as a limitation on the disclosure or the claims to follow.

**A. Introduction**

The traditional study of receptors has always proceeded from the a priori assumption (historically based) that the endogenous ligand must first be identified before discovery could proceed to find antagonists and other molecules that could affect the receptor. Even in cases where an antagonist might have been known first, the search immediately extended to looking for the endogenous ligand. This mode of thinking has persisted in receptor research even after the discovery of constitutively activated receptors. What has not been heretofore recognized is that it is the active state of the receptor that is most useful for discovering agonists, partial agonists, and inverse agonists of the receptor. For those diseases which result from an overly active receptor or an under-active receptor, what is desired in a therapeutic drug is a compound which acts to diminish the active state of a receptor or enhance the activity of the receptor, respectively, not necessarily a drug which is an antagonist to the endogenous ligand. This is because a compound that reduces or enhances the activity of the active receptor state need not bind at the same site as the endogenous ligand. Thus, as taught by a method of this invention, any search for therapeutic compounds should start by screening compounds against the ligand-independent active state.

Screening candidate compounds against non-endogenous, constitutively activated GPCRs allows for the direct identification of candidate compounds which act at these cell surface receptors, without requiring any prior knowledge or use of the receptor's endogenous ligand. By determining areas within the body where the endogenous version of such GPCRs are expressed and/or over-expressed, it is possible to determine related disease/disorder states which are associated with the expression and/or over-expression of these receptors; such an approach is disclosed in this patent document.

**B. Disease/Disorder Identification and/or Selection**

Most preferably, inverse agonists to the non-endogenous, constitutively activated GPCRs can be identified using the materials of this invention. Such inverse agonists are ideal candidates as lead compounds in drug discovery programs for treating diseases related to these receptors. Because of the ability to directly identify inverse agonists, partial agonists or agonists to these receptors, thereby allowing for the development of pharmaceutical compositions, a search, for diseases and disorders associated with these receptors is possible. For example, scanning both diseased and normal tissue samples for the presence of these receptor now becomes more than an academic exercise or one which might be pursued along the path of identifying, in the case of an orphan receptor, an endogenous ligand. Tissue scans can be conducted across a broad range of healthy and diseased tissues. Such tissue scans provide a preferred first step in associating a specific receptor with a disease and/or disorder.

Preferably, the DNA sequence of the endogenous GPCR is used to make a probe for either radiolabeled cDNA or RT-PCR identification of the expression of the GPCR in tissue samples. The presence of a receptor in a diseased tissue, or the presence of the receptor at elevated or decreased concentrations in diseased tissue compared to a normal tissue, can be preferably utilized to identify a correlation with that disease. Receptors can equally well be localized to regions of organs by this technique. Based on the known functions of the specific tissues to which the receptor is localized, the putative functional role of the receptor can be deduced.

**C. A "Human GPCR Proline Marker" Algorithm and the Creation of Non-Endogenous, Constitutively-Active Human GPCRs**

Among the many challenges facing the biotechnology arts is the unpredictability in gleaning genetic information from one species and correlating that information to another species

– nowhere in this art does this problem evidence more annoying exacerbation than in the genetic sequences that encode nucleic acids and proteins. Thus, for consistency and because of the highly unpredictable nature of this art, the following invention is limited, in terms of mammals, to human GPCRs – applicability of this invention to other mammalian species, while a potential possibility, is considered beyond mere rote application.

In general, when attempting to apply common "rules" from one related protein sequence to another or from one species to another, the art has typically resorted to sequence alignment, *i.e.*, sequences are linearized and attempts are then made to find regions of commonality between two or more sequences. While useful, this approach does not always prove to result in meaningful information. In the case of GPCRs, while the general structural motif is identical for all GPCRs, the variations in lengths of the TMs, ECs and ICs make such alignment approaches from one GPCR to another difficult at best. Thus, while it may be desirable to apply a consistent approach to, *e.g.*, constitutive activation from one GPCR to another, because of the great diversity in sequence length, fidelity, etc from one GPCR to the next, a generally applicable, and readily successful mutational alignment approach is in essence not possible. In an analogy, such an approach is akin to having a traveler start a journey at point A by giving the traveler dozens of different maps to point B, without any scale or distance markers on any of the maps, and then asking the traveler to find the shortest and most efficient route to destination B only by using the maps. In such a situation, the task can be readily simplified by having (a) a common "place-marker" on each map, and (b) the ability to measure the distance from the place-marker to destination B – this, then, will allow the traveler to select the most efficient from starting-point A to destination B.

In essence, a feature of the invention is to provide such coordinates within human GPCRs

that readily allows for creation of a constitutively active form of the human GPCRs.

As those in the art appreciate, the transmembrane region of a cell is highly hydrophobic; thus, using standard hydrophobicity plotting techniques, those in the art are readily able to determine the TM regions of a GPCR, and specifically TM6 (this same approach is also applicable to determining the EC and IC regions of the GPCR). It has been discovered that within the TM6 region of human GPCRs, a common proline residue (generally near the middle of TM6), acts as a constitutive activation "marker." By counting 15 amino acids from the proline marker, the 16$^{th}$ amino acid (which is located in the IC3 loop), when mutated from its endogenous form to a non-endogenous form, leads to constitutive activation of the receptor. For convenience, we refer to this as the "Human GPCR Proline Marker" Algorithm. Although the non-endogenous amino acid at this position can be any of the amino acids, most preferably, the non-endogenous amino acid is lysine. While not wishing to be bound by any theory, we believe that this position itself is unique and that the mutation at this location impacts the receptor to allow for constitutive activation.

We note that, for example, when the endogenous amino acid at the 16$^{th}$ position is already lysine (as is the case with GPR4 and GPR32), then in order for X to be a non-endogenous amino acid, it must be other than lysine; thus, in those situations where the endogenous GPCR has an endogenous lysine residue at the 16$^{th}$ position, the non-endogenous version of that GPCR preferably incorporates an amino acid other than lysine, preferably alanine, histidine and arginine, at this position. Of further note, it has been determined that GPR4 appears to be linked to Gs and active in its endogenous form (data not shown).

Because there are only 20 naturally occurring amino acids (although the use of non-naturally occurring amino acids is also viable), selection of a particular non-endogenous amino

acid for substitution at this 16th position is viable and allows for efficient selection of a non-endogenous amino acid that fits the needs of the investigator. However, as noted, the more preferred non-endogenous amino acids at the 16th position are lysine, hisitidine, arginine and alanine, with lysine being most preferred. Those of ordinary skill in the art are credited with the ability to readily determine proficient methods for changing the sequence of a codon to achieve a desired mutation.

It has also been discovered that occasionally, but not always, the proline residue marker will be preceded in TM6 by W2 (*i.e.*, $W2P^{1}AA_{1,5}X$) where W is tryptophan and 2 is any amino acid residue.

Our discovery, amongst other things, negates the need for unpredictable and complicated sequence alignment approaches commonly used by the art. Indeed, the strength of our discovery, while an algorithm in nature, is that it can be applied in a facile manner to human GPCRs, with dexterous simplicity by those in the art, to achieve a unique and highly useful end-product, *i.e.*, a constitutively activated version of a human GPCR. Because many years and significant amounts of money will be required to determine the endogenous ligands for the human GPCRs that the Human Genome project is uncovering, the disclosed invention not only reduces the time necessary to positively exploit this sequence information, but at significant cost-savings. This approach truly validates the importance of the Human Genome Project because it allows for the utilization of genetic information to not only understand the role of the GPCRs in, *e.g.*, diseases, but also provides the opportunity to improve the human condition.

**D. Screening of Candidate Compounds**

**1. Generic GPCR screening assay techniques**

When a G protein receptor becomes constitutively active, it couples to a G protein (*e.g.*,

Gq, Gs, Gi, Go) and stimulates release and subsequent binding of GTP to the G protein. The G protein then acts as a GTPase and slowly hydrolyzes the GTP to GDP, whereby the receptor, under normal conditions, becomes deactivated. However, constitutively activated receptors, including the non-endogenous, human constitutively active GPCRs of the present invention, continue to exchange GDP for GTP. A non-hydrolyzable analog of GTP, [$^{35}$S]GTPγS, can be used to monitor enhanced binding to G proteins present on membranes which express constitutively activated receptors. It is reported that [$^{35}$S]GTPγS can be used to monitor G protein coupling to membranes in the absence and presence of ligand. An example of this monitoring, among other examples well-known and available to those in the art, was reported by Traynor and Nahorski in 1995. The preferred use of this assay system is for initial screening of candidate compounds because the system is generically applicable to all G protein-coupled receptors regardless of the particular G protein that interacts with the intracellular domain of the receptor.

**B 2. Specific GPCR screening assay techniques**

C Once candidate compounds are identified using the "generic" G protein-coupled receptor assay (*i.e.*, an assay to select compounds that are agonists, partial agonists, or inverse agonists), further screening to confirm that the compounds have interacted at the receptor site is preferred. For example, a compound identified by the "generic" assay may not bind to the receptor, but may instead merely "uncouple" the G protein from the intracellular domain.

***a. Gs and Gi.***

Gs stimulates the enzyme adenylyl cyclase. Gi (and Go), on the other hand, inhibit this enzyme. Adenylyl cyclase catalyzes the conversion of ATP to cAMP; thus,

constitutively activated GPCRs that couple the Gs protein are associated with increased cellular levels of cAMP. On the other hand, constitutively activated GPCRs that couple the Gi (or Go) protein are associated with decreased cellular levels of cAMP. *See, generally*, "Indirect Mechanisms of Synaptic Transmission," Chpt. 8, From Neuron To Brain (3rd Ed.) Nichols, J.G. et al eds. Sinauer Associates, Inc. (1992). Thus, assays that detect cAMP can be utilized to determine if a candidate compound is, *e.g.*, an inverse agonist to the receptor (*i.e.*, such a compound would decrease the levels of cAMP). A variety of approaches known in the art for measuring cAMP can be utilized; a most preferred approach relies upon the use of anti-cAMP antibodies in an ELISA-based format. Another type of assay that can be utilized is a whole cell second messenger reporter system assay. Promoters on genes drive the expression of the proteins that a particular gene encodes. Cyclic AMP drives gene expression by promoting the binding of a cAMP-responsive DNA binding protein or transcription factor (CREB) which then binds to the promoter at specific sites called cAMP response elements and drives the expression of the gene. Reporter systems can be constructed which have a promoter containing multiple cAMP response elements before the reporter gene, *e.g.*, β-galactosidase or luciferase. Thus, a constitutively activated Gs-linked receptor causes the accumulation of cAMP that then activates the gene and expression of the reporter protein. The reporter protein such as β-galactosidase or luciferase can then be detected using standard biochemical assays (Chen et al. 1995). With respect to GPCRs that link to Gi (or Go), and thus decrease levels of cAMP, an approach to the screening of, *e.g.*, inverse agonists, based upon utilization of receptors that link to Gs (and thus increase levels of cAMP) is disclosed in the Example section with respect to GPR17 and GPR30.

*b. Go and Gq.*

Gq and Go are associated with activation of the enzyme phospholipase C, which in turn hydrolyzes the phospholipid $PIP_2$, releasing two intracellular messengers: diacycloglycerol (DAG) and inistol 1,4,5-triphoisphate ($IP_3$). Increased accumulation of $IP_3$ is associated with activation of Gq- and Go-associated receptors. *See, generally*, "Indirect Mechanisms of Synaptic Transmission," Chpt. 8, From Neuron To Brain (3rd Ed.) Nichols, J.G. et al eds. Sinauer Associates, Inc. (1992). Assays that detect $IP_3$ accumulation can be utilized to determine if a candidate compound is, *e.g.*, an inverse agonist to a Gq- or Go-associated receptor (*i.e.*, such a compound would decrease the levels of $IP_3$). Gq-associated receptors can also been examined using an AP1 reporter assay in that Gq-dependent phospholipase C causes activation of genes containing AP1 elements; thus, activated Gq-associated receptors will evidence an increase in the expression of such genes, whereby inverse agonists thereto will evidence a decrease in such expression, and agonists will evidence an increase in such expression. Commercially available assays for such detection are available.

**E. Medicinal Chemistry**

Generally, but not always, direct identification of candidate compounds is preferably conducted in conjunction with compounds generated via combinatorial chemistry techniques, whereby thousands of compounds are randomly prepared for such analysis. Generally, the results of such screening will be compounds having unique core structures; thereafter, these compounds are preferably subjected to additional chemical modification around a preferred core structure(s) to further enhance the medicinal properties thereof. Such techniques are

known to those in the art and will not be addressed in detail in this patent document.

**F. Pharmaceutical Compositions**

Candidate compounds selected for further development can be formulated into pharmaceutical compositions using techniques well known to those in the art. Suitable pharmaceutically-acceptable carriers are available to those in the art; for example, *see* Remington's Pharmaceutical Sciences, 16th Edition, 1980, Mack Publishing Co., (Oslo et al., eds.)

**G. Other Utility**

Although a preferred use of the non-endogenous versions of the disclosed human GPCRs is for the direct identification of candidate compounds as inverse agonists, agonists or partial agonists (preferably for use as pharmaceutical agents), these receptors can also be utilized in research settings. For example, in vitro and in vivo systems incorporating these receptors can be utilized to further elucidate and understand the roles of the receptors in the human condition, both normal and diseased, as well understanding the role of constitutive activation as it applies to understanding the signaling cascade. A value in these non-endogenous receptors is that their utility as a research tool is enhanced in that, because of their unique features, the disclosed receptors can be used to understand the role of a particular receptor in the human body before the endogenous ligand therefor is identified. Other uses of the disclosed receptors will become apparent to those in the art based upon, *inter alia*, a review of this patent document.

## EXAMPLES

The following examples are presented for purposes of elucidation, and not limitation, of the present invention. Following the teaching of this patent document that a mutational cassette may be utilized in the IC3 loop of human GPCRs based upon a position relative to a proline residue in TM6 to constitutively activate the receptor, and while specific nucleic acid

and amino acid sequences are disclosed herein, those of ordinary skill in the art are credited with the ability to make minor modifications to these sequences while achieving the same or substantially similar results reported below. Particular approaches to sequence mutations are within the purview of the artisan based upon the particular needs of the artisan.

**Example 1**
**Preparation of Endogenous Human GPCRs**

A variety of GPCRs were utilized in the Examples to follow. Some endogenous human GPCRs were graciously provided in expression vectors (as acknowledged below) and other endogenous human GPCRs were synthesized *de novo* using publicly-available sequence information.

**1. GPR1 (GenBank Accession Number: U13666)**

The human cDNA sequence for GPR1 was provided in pRcCMV by Brian O'Dowd (University of Toronto). GPR1 cDNA (1.4kB fragment) was excised from the pRcCMV vector as a NdeI-XbaI fragment and was subcloned into the NdeI-XbaI site of pCMV vector (*see* Figure 3). Nucleic acid (SEQ.ID.NO.: 1) and amino acid (SEQ.ID.NO.: 2) sequences for human GPR1 were thereafter determined and verified.

**2. GPR4 (GenBank Accession Numbers: L36148, U35399, U21051)**

The human cDNA sequence for GPR4 was provided in pRcCMV by Brian O'Dowd (University of Toronto). GPR1 cDNA (1.4kB fragment) was excised from the pRcCMV vector as an ApaI(blunted)-XbaI fragment and was subcloned (with most of the 5' untranslated region removed) into HindIII(blunted)-XbaI site of pCMV vector. Nucleic acid (SEQ.ID.NO.: 3) and amino acid (SEQ.ID.NO.: 4) sequences for human GPR4 were thereafter determined and verified.

**3. GPR5 (GenBank Accession Number: L36149)**

The cDNA for human GPR5 was generated and cloned into pCMV expression vector as follows: PCR was performed using genomic DNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each of the 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min; 64°C for 1min; and 72°C for 1.5 min. The 5' PCR primer contained an EcoRI site with the sequence:

5'-TATGAATTCAGATGCTCTAAACGTCCCTGC-3' (SEQ.ID.NO.: 5)

and the 3' primer contained BamHI site with the sequence:

5'-TCCGGATCCACCTGCACCTGCGCCTGCACC-3' (SEQ.ID.NO.: 6).

The 1.1 kb PCR fragment was digested with EcoRI and BamHI and cloned into EcoRI-BamHI site of PCMV expression vector. Nucleic acid (SEQ.ID.NO.: 7) and amino acid (SEQ.ID.NO.: 8) sequences for human GPR5 were thereafter determined and verified.

**4. GPR7 (GenBank Accession Number: U22491)**

The cDNA for human GPR7 was generated and cloned into pCMV expression vector as follows: PCR condition- PCR was performed using genomic DNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each of the 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min; 62°C for 1min; and 72°C for 1min and 20 sec. The 5' PCR primer contained a HindIII site with the sequence:

5'-GCAAGCTTGGGGGACGCCAGGTCGCCGGCT-3' (SEQ.ID.NO.: 9)

and the 3' primer contained a BamHI site with the sequence:

5'-GCGGATCCGGACGCTGGGGGAGTCAGGCTGC-3' (SEQ.ID.NO.: 10).

The 1.1 kb PCR fragment was digested with HindIII and BamHI and cloned into HindIII-BamHI site of pCMV expression vector. Nucleic acid (SEQ.ID.NO.: 11) and amino acid (SEQ.ID.NO.:

12) sequences for human GPR7 were thereafter determined and verified.

**5. GPR8 (GenBank Accession Number: U22492)**

The cDNA for human GPR8 was generated and cloned into pCMV expression vector as follows: PCR was performed using genomic DNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each of the 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min; 62°C for 1min; and 72 °C for 1min and 20 sec. The 5' PCR primer contained an EcoRI site with the sequence:

5'-CGGAATTCGTCAACGGTCCCAGCTACAATG-3' (SEQ.ID.NO.: 13).

and the 3' primer contained a BamHI site with the sequence:

5'-ATGGATCCCAGGCCCTTCAGCACCGCAATAT-3'(SEQ.ID.NO.: 14).

The 1.1 kb PCR fragment was digested with EcoRI and BamHI and cloned into EcoRI-BamHI site of PCMV expression vector. All 4 cDNA clones sequenced contained a possible polymorphism involving a change of amino acid 206 from Arg to Gln. Aside from this difference, nucleic acid (SEQ.ID.NO.: 15) and amino acid (SEQ.ID.NO.: 16) sequences for human GPR8 were thereafter determined and verified.

**6. GPR9 (GenBank Accession Number: X95876)**

The cDNA for human GPR9 was generated and cloned into pCMV expression vector as follows: PCR was performed using a clone (provided by Brian O'Dowd) as template and pfu polymerase (Stratagene) with the buffer system provided by the manufacturer supplemented with 10% DMSO, 0.25 μM of each primer, and 0.5 mM of each of the 4 nucleotides. The cycle condition was 25 cycles of: 94°C for 1 min; 56°C for 1min; and 72 °C for 2.5 min. The 5' PCR primer contained an EcoRI site with the sequence:

5'-ACGAATTCAGCCATGGTCCTTGAGGTGAGTGACCACCAAGTGCTAAAT-3' (SEQ.ID.NO.: 17)

and the 3' primer contained a BamHI site with the sequence:

5'-GAGGATCCTGGAATGCGGGGAAGTCAG-3' (SEQ.ID.NO.: 18).

The 1.2 kb PCR fragment was digested with EcoRI and cloned into EcoRI-SmaI site of PCMV expression vector. Nucleic acid (SEQ.ID.NO.: 19) and amino acid (SEQ.ID.NO.: 20) sequences for human GPR9 were thereafter determined and verified.

**7. GPR9-6 (GenBank Accession Number: U45982)**

The cDNA for human GPR9-6 was generated and cloned into pCMV expression vector as follows: PCR was performed using genomic DNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each of the 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min; 62°C for 1min; and 72 °C for 1 min and 20 sec. The 5' PCR primer was kinased with the sequence:

5'-TTAAGCTTGACCTAATGCCATCTTGTGTCC-3' (SEQ.ID.NO.: 21)

and the 3' primer contained a BamHI site with the sequence:

5'-TTGGATCCAAAAGAACCATGCACCTCAGAG-3' (SEQ.ID.NO.: 22).

The 1.2 kb PCR fragment was digested with BamHI and cloned into EcoRV-BamHI site of pCMV expression vector. Nucleic acid (SEQ.ID.NO.: 23) and amino acid (SEQ.ID.NO.: 24) sequences for human GPR9-6 were thereafter determined and verified.

**8. GPR10 (GenBank Accession Number: U32672)**

The human cDNA sequence for GPR10 was provided in pRcCMV by Brian O'Dowd (University of Toronto). GPR10 cDNA (1.3kB fragment) was excised from the pRcCMV vector as an EcoRI-XbaI fragment and was subcloned into EcoRI-XbaI site of pCMV

vector. Nucleic acid (SEQ.ID.NO.: 25) and amino acid (SEQ.ID.NO.: 26) sequences for human GPR10 were thereafter determined and verified.

**9. GPR15 (GenBank Accession Number: U34806)**

The human cDNA sequence for GPR15 was provided in pCDNA3 by Brian O'Dowd (University of Toronto). GPR15 cDNA (1.5kB fragment) was excised from the pCDNA3 vector as a HindIII-Bam fragment and was subcloned into HindIII-Bam site of pCMV vector. Nucleic acid (SEQ.ID.NO.: 27) and amino acid (SEQ.ID.NO.: 28) sequences for human GPR15 were thereafter determined and verified.

**10. GPR17 (GenBank Accession Number: Z94154)**

The cDNA for human GPR17 was generated and cloned into pCMV expression vector as follows: PCR was performed using genomic DNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min; 56°C for 1min and 72 °C for 1 min and 20 sec. The 5' PCR primer contained an EcoRI site with the sequence:

5'-CTAGAATTCTGACTCCAGCCAAAGCATGAAT-3' (SEQ.ID.NO.: 29)and the 3' primer contained a BamHI site with the sequence:

5'-GCTGGATCCTAAACAGTCTGCGCTCGGCCT-3' (SEQ.ID.NO.: 30).

The 1.1 kb PCR fragment was digested with EcoRI and BamHI and cloned into EcoRI-BamHI site of pCMV expression vector. Nucleic acid (SEQ.ID.NO.: 31) and amino acid (SEQ.ID.NO.: 32) sequences for human GPR17 were thereafter determined and verified.

**11. GPR18 (GenBank Accession Number: L42324)**

The cDNA for human GPR18 was generated and cloned into pCMV expression

vector as follows: PCR was performed using genomic DNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each of the 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min; 54°C for 1min; and 72 °C for 1min and 20 sec. The 5' PCR primer was kinased with the sequence:

5'-ATAAGATGATCACCCTGAACAATCAAGAT -3' (SEQ.ID.NO.: 33)

and the 3' primer contained an EcoRI site with the sequence:

5'-TCCGAATTCATAACATTTCACTGTTTATATTGC-3' (SEQ.ID.NO.: 34).

The 1.0 kb PCR fragment was digested with EcoRI and cloned into blunt-EcoRI site of pCMV expression vector. All 8 cDNA clones sequenced contained 4 possible polymorphisms involving changes of amino acid 12 from Thr to Pro, amino acid 86 from Ala to Glu, amino acid 97 from Ile to Leu and amino acid 310 from Leu to Met. Aside from these changes, nucleic acid (SEQ.ID.NO.: 35) and amino acid (SEQ.ID.NO.: 36) sequences for human GPR18 were thereafter determined and verified.

**12. GPR20 (GenBank Accession Number: U66579)**

The cDNA for human GPR20 was generated and cloned into pCMV expression vector as follows: PCR was performed using genomic DNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each of the 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min; 62°C for 1min; and 72 °C for 1 min and 20 sec. The 5' PCR primer was kinased with the sequence:

5'-CCAAGCTTCCAGGCCTGGGGTGTGCTGG-3' (SEQ.ID.NO.: 37)

and the 3' primer contained a BamHI site with the sequence:

5'-ATGGATCCTGACCTTCGGCCCCTGGCAGA-3' (SEQ.ID.NO.: 38).

The 1.2 kb PCR fragment was digested with BamHI and cloned into EcoRV-BamHI site of

PCMV expression vector. Nucleic acid (SEQ.ID.NO.: 39) and amino acid (SEQ.ID.NO.: 40) sequences for human GPR20 were thereafter determined and verified.

**13. GPR21 (GenBank Accession Number: U66580)**

The cDNA for human GPR21 was generated and cloned into pCMV expression vector as follows: PCR was performed using genomic DNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each of the 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min; 62°C for 1min; and 72 °C for 1 min and 20 sec. The 5' PCR primer was kinased with the sequence:

5'-GAGAATTCACTCCTGAGCTCAAGATGAACT-3' (SEQ.ID.NO.: 41)

and the 3' primer contained a BamHI site with the sequence:

5'-CGGGATCCCCGTAACTGAGCCACTTCAGAT-3' (SEQ.ID.NO.: 42).

The 1.1 kb PCR fragment was digested with BamHI and cloned into EcoRV-BamHI site of pCMV expression vector. Nucleic acid (SEQ.ID.NO.: 43) and amino acid (SEQ.ID.NO.: 44) sequences for human GPR21 were thereafter determined and verified.

**14. GPR22 (GenBank Accession Number: U66581)**

The cDNA for human GPR22 was generated and cloned into pCMV expression vector as follows: PCR was performed using genomic DNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each of the 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min; 50°C for 1min; and 72 °C for 1.5 min. The 5' PCR primer was kinased with the sequence:

5'-TCCCCCGGGAAAAAAACCAACTGCTCCAAA-3' (SEQ.ID.NO.: 45)

and the 3' primer contained a BamHI site with the sequence:

5'-TAGGATCCATTTGAATGTGGATTTGGTGAAA-3' (SEQ.ID.NO.: 46).

The 1.38 kb PCR fragment was digested with BamHI and cloned into EcoRV-BamHI site of pCMV expression vector. Nucleic acid (SEQ.ID.NO.: 47) and amino acid (SEQ.ID.NO.: 48) sequences for human GPR22 were thereafter determined and verified.

**15. GPR24 (GenBank Accession Number: U71092)**

The cDNA for human GPR24 was generated and cloned into pCMV expression vector as follows: PCR was performed using genomic DNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min; 56°C for 1min; and 72 °C for 1 min and 20 sec. The 5' PCR primer contains a HindIII site with the sequence:

5'-GTGAAGCTTGCCTCTGGTGCCTGCAGGAGG-3' (SEQ.ID.NO.: 49)

and the 3' primer contains an EcoRI site with the sequence:

5'-GCAGAATTCCCGGTGGCGTGTTGTGGTGCCC-3' (SEQ.ID.NO.: 50).

The 1.3 kb PCR fragment was digested with HindIII and EcoRI and cloned into HindIII-EcoRI site of pCMV expression vector. The nucleic acid (SEQ.ID.NO.: 51) and amino acid sequence (SEQ.ID.NO.: 52) for human GPR24 were thereafter determined and verified.

**16. GPR30 (GenBank Accession Number: U63917)**

The cDNA for human GPR30 was generated and cloned as follows: the coding sequence of GPR30 (1128bp in length) was amplified from genomic DNA using the primers:

5'-GGCGGATCCATGGATGTGACTTCCCAA-3' (SEQ.ID.NO.: 53) and

5'-GGCGGATCCCTACACGGCACTGCTGAA-3' (SEQ.ID.NO.: 54).

The amplified product was then cloned into a commercially available vector, pCR2.1 (Invitrogen),

using a "TOPO-TA Cloning Kit" (Invitrogen, #K4500-01), following manufacturer instructions. The full-length GPR30 insert was liberated by digestion with BamH1, separated from the vector by agarose gel electrophoresis, and purified using a Sephaglas Bandprep™ Kit (Pharmacia, # 27-9285-01) following manufacturer instructions. The nucleic acid (SEQ.ID.NO.: 55) and amino acid sequence (SEQ.ID.NO.: 56) for human GPR30 were thereafter determined and verified.

**17. GPR31 (GenBank Accession Number: U65402)**

The cDNA for human GPR31 was generated and cloned into pCMV expression vector as follows: PCR was performed using genomic DNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each of the 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min; 58°C for 1min; and 72 °C for 2 min. The 5' PCR primer contained an EcoRI site with the sequence:

5'-AAGGAATTCACGGCCGGGTGATGCCATTCCC-3' (SEQ.ID.NO.: 57)

and the 3' primer contained a BamHI site with the sequence:

5'-GGTGGATCCATAAACACGGGCGTTGAGGAC -3' (SEQ.ID.NO.: 58).

The 1.0 kb PCR fragment was digested with EcoRI and BamHI and cloned into EcoRI-BamHI site of pCMV expression vector. Nucleic acid (SEQ.ID.NO.: 59) and amino acid (SEQ.ID.NO.: 60) sequences for human GPR31 were thereafter determined and verified.

**18. GPR32 (GenBank Accession Number: AF045764)**

The cDNA for human GPR32 was generated and cloned into pCMV expression vector as follows: PCR was performed using genomic DNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min; 56°C for 1min; and 72 °C for 1 min and 20 sec. The 5' PCR primer contained an EcoRI site with the

sequence:

5'-TAAGAATTCCATAAAAATTATGGAATGG-3' (SEQ.ID.NO.:243)

and the 3' primer contained a BamHI site with the sequence:

5'-CCAGGATCCAGCTGAAGTCTTCCATCATTC-3' (SEQ.ID.NO.: 244).

The 1.1 kb PCR fragment was digested with EcoRI and BamHI and cloned into EcoRI-BamHI site of pCMV expression vector. Nucleic acid (SEQ.ID.NO.: 245) and amino acid (SEQ.ID.NO.: 246) sequences for human GPR32 were thereafter determined and verified.

**19. GPR40 (GenBank Accession Number: AF024687)**

The cDNA for human GPR40 was generated and cloned into pCMV expression vector as follows: PCR was performed using genomic DNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min, 65°C for 1min and 72 °C for 1 min and 10 sec. The 5' PCR primer contained an EcoRI site with the sequence

5'-GCAGAATTCGGCGGCCCCATGGACCTGCCCCC-3' (SEQ.ID.NO.: 247)

and the 3' primer contained a BamHI site with the sequence

5'-GCTGGATCCCCCGAGCAGTGGCGTTACTTC-3' (SEQ.ID.NO.: 248).

The 1 kb PCR fragment was digested with EcoRI and BamHI and cloned into EcoRI-BamHI site of pCMV expression vector. Nucleic acid (SEQ.ID.NO.: 249) and amino acid (SEQ.ID.NO.: 250) sequences for human GPR40 were thereafter determined and verified.

**20. GPR41 (GenBank Accession Number AF024688)**

The cDNA for human GPR41 was generated and cloned into pCMV expression vector as follows: PCR was performed using genomic DNA as template and rTth polymerase

(Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each 4 nucleotides. The cycle condition was 30 cycles of 94°C for 1 min, 65°C for 1min and 72 °C for 1 min and 10 sec. The 5' PCR primer contained an HindIII site with the sequence:

5'-CTCAAGCTTACTCTCTCTCACCAGTGGCCAC-3' (SEQ.ID.NO.: 251)

and the 3' primer was kinased with the sequence

5'-CCCTCCTCCCCCGGAGGACCTAGC-3' (SEQ.ID.NO.: 252).

The 1 kb PCR fragment was digested with HindIII and cloned into HindIII-blunt site of pCMV expression vector. Nucleic acid (SEQ.ID.NO.: 253) and amino acid (SEQ.ID.NO.: 254) sequences for human GPR41 were thereafter determined and verified.

**21. GPR43 (GenBank Accession Number AF024690)**

The cDNA for human GPR43 was generated and cloned into pCMV expression vector as follows: PCR was performed using genomic DNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min; 65°C for 1min; and 72 °C for 1 min and 10 sec. The 5' PCR primer contains an HindIII site with the sequence:

5'-TTTAAGCTTCCCCTCCAGGATGCTGCCGGAC-3' (SEQ.ID.NO.: 255)

and the 3' primer contained an EcoRI site with the sequence:

5'-GGCGAATTCTGAAGGTCCAGGGAAACTGCTA-3' (SEQ.ID.NO. 256).

The 1 kb PCR fragment was digested with HindIII and EcoRI and cloned into HindIII-EcoRI site of pCMV expression vector. Nucleic acid (SEQ.ID.NO.: 257) and amino acid (SEQ.ID.NO.: 258) sequences for human GPR43 were thereafter determined and verified.

**22. APJ (GenBank Accession Number: U03642)**

Human APJ cDNA (in pRcCMV vector) was provided by Brian O'Dowd (University of Toronto). The human APJ cDNA was excised from the pRcCMV vector as an EcoRI-XbaI (blunted) fragment and was subcloned into EcoRI-SmaI site of pCMV vector. Nucleic acid (SEQ.ID.NO.: 61) and amino acid (SEQ.ID.NO.: 62) sequences for human APJ were thereafter determined and verified.

**23. BLR1 (GenBank Accession Number: X68149)**

The cDNA for human BLR1 was generated and cloned into pCMV expression vector as follows: PCR was performed using thymus cDNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each of the 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min; 62°C for 1min; and 72 °C for 1 min and 20 sec. The 5' PCR primer contained an EcoRI site with the sequence:

5'-TGAGAATTCTGGTGACTCACAGCCGGCACAG-3' (SEQ.ID.NO.: 63):

and the 3' primer contained a BamHI site with the sequence:

5'-GCCGGATCCAAGGAAAAGCAGCAATAAAAGG-3' (SEQ.ID.NO.: 64). The 1.2 kb PCR fragment was digested with EcoRI and BamHI and cloned into EcoRI-BamHI site of pCMV expression vector. Nucleic acid (SEQ.ID.NO.: 65) and amino acid (SEQ.ID.NO.: 66) sequences for human BLR1 were thereafter determined and verified.

**24. CEPR (GenBank Accession Number: U77827)**

The cDNA for human CEPR was generated and cloned into pCMV expression vector as follows: PCR was performed using genomic DNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and

0.2 mM of each of the 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min; 65°C for 1min; and 72 °C for 1 min and 20 sec. The 5' PCR primer was kinased with the sequence:

5'-CAAAGCTTGAAAGCTGCACGGTGCAGAGAC-3' (SEQ.ID.NO.:67)

and the 3' primer contained a BamHI site with the sequence:

5'-GCGGATCCCGAGTCACACCCTGGCTGGGCC-3' (SEQ.ID.NO.: 68).

The 1.2 kb PCR fragment was digested with BamHI and cloned into EcoRV-BamHI site of pCMV expression vector. Nucleic acid (SEQ.ID.NO.: 69) and amino acid (SEQ.ID.NO.: 70) sequences for human CEPR were thereafter determined and verified.

**25. EBI1 (GenBank Accession Number: L31581)**

The cDNA for human EBI1 was generated and cloned into pCMV expression vector as follows: PCR was performed using thymus cDNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each of the 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min; 62°C for 1min; and 72 °C for 1 min and 20 sec. The 5' PCR primer contained an EcoRI site with the sequence:

5'-ACAGAATTCCTGTGTGGTTTTACCGCCCAG-3' (SEQ.ID.NO.: 71)

and the 3' primer contained a BamHI site with the sequence:

5'-CTCGGATCCAGGCAGAAGAGTCGCCTATGG-3' (SEQ.ID.NO.: 72).

The 1.2 kb PCR fragment was digested with EcoRI and BamHI and cloned into EcoRI-BamHI site of PCMV expression vector. Nucleic acid (SEQ.ID.NO.: 73) and amino acid (SEQ.ID.NO.: 74) sequences for human EBI1 were thereafter determined and verified.

**26. EBI2 (GenBank Accession Number: L08177)**

The cDNA for human EBI2 was generated and cloned into pCMV expression

vector as follows: PCR was performed using cDNA clone (graciously provided by Kevin Lynch, University of Virginia Health Sciences Center; the vector utilized was not identified by the source) as template and pfu polymerase (Stratagene) with the buffer system provided by the manufacturer supplemented with 10% DMSO, 0.25 μM of each primer, and 0.5 mM of each of the 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min; 60°C for 1min; and 72 °C for 1 min and 20 sec. The 5' PCR primer contained an EcoRI site with the sequence:

5'-CTGGAATTCACCTGGACCACCACCAATGGATA-3' (SEQ.ID.NO.: 75)

and the 3' primer contained a BamHI site with the sequence

5'-CTCGGATCCTGCAAAGTTTGTCATACAG TT-3' (SEQ.ID.NO.: 76).

The 1.2 kb PCR fragment was digested with EcoRI and BamHI and cloned into EcoRI-BamHI site of pCMV expression vector. Nucleic acid (SEQ.ID.NO.: 77) and amino acid (SEQ.ID.NO.: 78) sequences for human EBI2 were thereafter determined and verified.

**27. ETBR-LP2 (GenBank Accession Number: D38449)**

The cDNA for human ETBR-LP2 was generated and cloned into pCMV expression vector as follows: PCR was performed using brain cDNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each of the 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min; 65°C for 1min; and 72 °C for 1.5 min. The 5' PCR contained an EcoRI site with the sequence:

5'-CTGGAATTCTCCTGCTCATCCAGCCATGCGG -3' (SEQ.ID.NO.: 79)

and the 3' primer contained a BamHI site with the sequence:

5'-CCTGGATCCCCACCCCTACTGGGGCCTCAG -3' (SEQ.ID.NO.: 80).

The 1.5 kb PCR fragment was digested with EcoRI and BamHI and cloned into EcoRI-BamHI

site of pCMV expression vector. Nucleic acid (SEQ.ID.NO.: 81) and amino acid (SEQ.ID.NO.: 82) sequences for human ETBR-LP2 were thereafter determined and verified.

**28. GHSR (GenBank Accession Number: U60179)**

The cDNA for human GHSR was generated and cloned into pCMV expression vector as follows: PCR was performed using hippocampus cDNA as template and TaqPlus Precision polymerase (Stratagene) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each of the 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min; 68°C for 1min; and 72°C for 1 min and 10 sec. For first round PCR, the 5' PCR primer sequence was:

5'-ATGTGGAACGCGACGCCCAGCG-3' (SEQ.ID.NO.: 83)

and the 3' primer sequence was:

5'-TCATGTATTAATACTAGATTCT-3' (SEQ.ID.NO.: 84).

Two microliters of the first round PCR was used as template for the second round PCR where the 5' primer was kinased with sequence:

5'-TACCATGTGGAACGCGACGCCCAGCGAAGAGCCGGGGT-3'(SEQ.ID.NO.:85)

and the 3' primer contained an EcoRI site with the sequence:

5'-CGGAATTCATGTATTAATACTAGATTCTGTCCAGGCCCG-3'(SEQ.ID.NO.:86).

The 1.1 kb PCR fragment was digested with EcoRI and cloned into blunt-EcoRI site of pCMV expression vector. Nucleic acid (SEQ.ID.NO.: 87) and amino acid (SEQ.ID.NO.: 88) sequences for human GHSR were thereafter determined and verified.

**29. GPCR-CNS (GenBank Accession Number: AFO17262)**

The cDNA for human GPCR-CNS was generated and cloned into pCMV expression vector as follows: PCR was performed using brain cDNA as template and rTth

polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each of the 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min; 65°C for 1min; and 72°C for 2 min. The 5' PCR primer contained a HindIII site with the sequence:

5'-GCAAGCTTGTGCCCTCACCAAGCCATGCGAGCC-3' (SEQ.ID.NO.: 89)

and the 3' primer contained an EcoRI site with the sequence:

5'-CGGAATTCAGCAATGAGTTCCGACAGAAGC-3' (SEQ.ID.NO.: 90).

The 1.9 kb PCR fragment was digested with HindIII and EcoRI and cloned into HindIII-EcoRI site of pCMV expression vector. All nine clones sequenced contained a potential polymorphism involving a S284C change. Aside from this difference, nucleic acid (SEQ.ID.NO.: 91) and amino acid (SEQ.ID.NO.: 92) sequences for human GPCR-CNS were thereafter determined and verified.

**30. GPR-NGA (GenBank Accession Number: U55312)**

The cDNA for human GPR-NGA was generated and cloned into pCMV expression vector as follows: PCR was performed using genomic DNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each of the 4 nucleotides. The cycle condition was 30 cycles of 94°C for 1 min, 56°C for 1min and 72°C for 1.5 min. The 5' PCR primer contained an EcoRI site with the sequence:

5'-CAGAATTCAGAGAAAAAAAGTGAATATGGTTTTT-3' (SEQ.ID.NO.: 93)

and the 3' primer contained a BamHI site with the sequence:

5'-TTGGATCCCTGGTGCATAACAATTGAAAGAAT-3' (SEQ.ID.NO.: 94).

The 1.3 kb PCR fragment was digested with EcoRI and BamHI and cloned into EcoRI-BamHI site of pCMV expression vector. Nucleic acid (SEQ.ID.NO.: 95) and amino acid (SEQ.ID.NO.:

96) sequences for human GPR-NGA were thereafter determined and verified.

**31. H9 (GenBank Accession Number: U52219)**

The cDNA for human HB954 was generated and cloned into pCMV expression vector as follows: PCR was performed using pituitary cDNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min, 62°C for 1min and 72°C for 2 min. The 5' PCR primer contains a HindIII site with the sequence:

5'-GGAAAGCTTAACGATCCCCAGGAGCAACAT-3' (SEQ.ID.NO.: 97)

and the 3' primer contains a BamHI site with the sequence:

5'-CTGGGATCCTACGAGAGCATTTTTCACACAG-3' (SEQ.ID.NO.: 98).

The 1.9 kb PCR fragment was digested with HindIII and BamHI and cloned into HindIII-BamHI site of pCMV expression vector. When compared to the published sequences, a different isoform with 12 bp in frame insertion in the cytoplasmic tail was also identified and designated "H9b." Both isoforms contain two potential polymorphisms involving changes of amino acid P320S and amino acid G448A. Isoform H9a contained another potential polymorphism of amino acid S493N, while isoform H9b contained two additional potential polymorphisms involving changes of amino acid I502T and amino acid A532T (corresponding to amino acid 528 of isoform H9a). Nucleic acid (SEQ.ID.NO.: 99) and amino acid (SEQ.ID.NO.: 100) sequences for human H9 were thereafter determined and verified (in the section below, both isoforms were mutated in accordance with the Human GPCR Proline Marker Algorithm).

**32. HB954 (GenBank Accession Number: D38449)**

The cDNA for human HB954 was generated and cloned into pCMV expression

vector as follows: PCR was performed using brain cDNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each of the 4 nucleotides. The cycle condition was 30 cycles of 94°C for 1 min, 58°C for 1min and 72 °C for 2 min. The 5' PCR contained a HindIII site with the sequence:

5'-TCCAAGCTTCGCCATGGGACATAACGGGAGCT -3' (SEQ.ID.NO.: 101)

and the 3' primer contained an EcoRI site with the sequence:

5'-CGTGAATTCCAAGAATTTACAATCCTTGCT -3' (SEQ.ID.NO.: 102).

The 1.6 kb PCR fragment was digested with HindIII and EcoRI and cloned into HindIII-EcoRI site of pCMV expression vector. Nucleic acid (SEQ.ID.NO.: 103) and amino acid (SEQ.ID.NO.: 104) sequences for human HB954 were thereafter determined and verified.

**33. HG38 (GenBank Accession Number: AF062006)**

The cDNA for human HG38 was generated and cloned into pCMV expression vector as follows: PCR was performed using brain cDNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each 4 nucleotides. The cycle condition was 30 cycles of 94°C for 1 min, 56°C for 1min and 72 °C for 1 min and 30 sec. Two PCR reactions were performed to separately obtain the 5' and 3' fragment. For the 5' fragment, the 5' PCR primer contained an HindIII site with the sequence:

5'-CCCAAGCTTCGGGCACCATGGACACCTCCC-3' (SEQ.ID.NO.: 259)

and the 3' primer contained a BamHIsite with the sequence:

5'-ACAGGATCCAAATGCACAGCACTGGTAAGC-3' (SEQ.ID.NO.: 260).

This 5' 1.5 kb PCR fragment was digested with HindIII and BamHI and cloned into an HindIII-BamHI site of pCMV. For the 3' fragment, the 5' PCR primer was kinased with the sequence:

5'-CTATAACTGGGTTACATGGTTTAAC-3' (SEQ.ID.NO. 261)

and the 3' primer contained an EcoRI site with the sequence:

5'-TTTGAATTCACATATTAATTAGAGACATGG-3' (SEQ.ID.NO.: 262).

The 1.4 kb 3' PCR fragment was digested with EcoRI and subcloned into a blunt-EcoRI site of pCMV vector. The 5' and 3' fragments were then ligated together through a common EcoRV site to generate the full length cDNA clone. Nucleic acid (SEQ.ID.NO.: 263) and amino acid (SEQ.ID.NO.: 264) sequences for human HG38 were thereafter determined and verified.

**34. HM74 (GenBank Accession Number: D10923)**

The cDNA for human HM74 was generated and cloned into pCMV expression vector as follows: PCR was performed using either genomic DNA or thymus cDNA (pooled) as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each of the 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min; 65°C for 1min; and 72°C for 1 min and 20 sec. The 5' PCR primer contained an EcoRI site with the sequence:

5'-GGAGAATTCACTAGGCGAGGCGCTCCATC-3' (SEQ.ID.NO.: 105)

and the 3' primer was kinased with the sequence:

5'-GGAGGATCCAGGAAACCTTAGGCCGAGTCC-3' (SEQ.ID.NO.:106).

The 1.3 kb PCR fragment was digested with EcoRI and cloned into EcoRI-SmaI site of pCMV expression vector. Clones sequenced revealed a potential polymorphism involving a N94K change. Aside from this difference, nucleic acid (SEQ.ID.NO.: 107) and amino acid (SEQ.ID.NO.: 108) sequences for human HM74 were thereafter determined and verified.

**35. MIG (GenBank Accession Numbers: AFO44600 and AFO44601)**

The cDNA for human MIG was generated and cloned into pCMV expression vector as follows: PCR was performed using genomic DNA as template and TaqPlus Precision

polymerase (Stratagene) for first round PCR or pfu polymerase (Stratagene) for second round PCR with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM (TaqPlus Precision) or 0.5 mM (pfu) of each of the 4 nucleotides. When pfu was used, 10% DMSO was included in the buffer. The cycle condition was 30 cycles of: 94°C for 1 min; 65°C for 1min; and 72 °C for: (a) 1 min for first round PCR; and (b) 2 min for second round PCR. Because there is an intron in the coding region, two sets of primers were separately used to generate overlapping 5' and 3' fragments. The 5' fragment PCR primers were:

5'-ACCATGGCTTGCAATGGCAGTGCGGCCAGGGGGCACT-3' (external sense) (SEQ.ID.NO.: 109)

and

5'-CGACCAGGACAAACAGCATCTTGGTCACTTGTCTCCGGC-3'(internal antisense) (SEQ.ID.NO.: 110).

The 3' fragment PCR primers were:

5'-GACCAAGATGCTGTTTGTCCTGGTCGTGGTGTTTGGCAT-3' (internal sense) (SEQ.ID.NO.: 111) and

5'-CGGAATTCAGGATGGATCGGTCTCTTGCTGCGCCT-3' (external antisense with an EcoRI site) (SEQ.ID.NO.: 112).

The 5' and 3' fragments were ligated together by using the first round PCR as template and the kinased external sense primer and external antisense primer to perform second round PCR. The 1.2 kb PCR fragment was digested with EcoRI and cloned into the blunt-EcoRI site of pCMV expression vector. Nucleic acid (SEQ.ID.NO.: 113) and amino acid (SEQ.ID.NO.: 114) sequences for human MIG were thereafter determined and verified.

**36. OGR1 (GenBank Accession Number: U48405)**

The cDNA for human OGR1 was generated and cloned into pCMV expression vector as follows: PCR was performed using genomic DNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each of the 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min; 65°C for 1min; and 72 °C for 1 min and 20 sec. The 5' PCR primer was kinased with the sequence:

5'-GGAAGCTTCAGGCCCAAAGATGGGGAACAT-3' (SEQ.ID.NO.: 115):

and the 3' primer contained a BamHI site with the sequence:

5'-GTGGATCCACCCGCGGAGGACCCAGGCTAG -3' (SEQ.ID.NO.: 116).

The 1.1 kb PCR fragment was digested with BamHI and cloned into the EcoRV-BamHI site of pCMV expression vector. Nucleic acid (SEQ.ID.NO.: 117) and amino acid (SEQ.ID.NO.: 118) sequences for human OGR1 were thereafter determined and verified.

**37. Serotonin $5HT_{2A}$**

The cDNA encoding endogenous human $5HT_{2A}$ receptor was obtained by RT-PCR using human brain poly-$A^+$ RNA; a 5' primer from the 5' untranslated region with an Xho I restriction site:

5'-GACCTCGAGTCCTTCTACACCTCATC-3' (SEQ.ID.NO: 119)

and a 3' primer from the 3' untranslated region containing an Xba I site:

5'-TGCTCTAGATTCCAGATAGGTGAAAACTTG-3' (SEQ.ID.NO: 120)

PCR was performed using either TaqPlus™ precision polymerase (Stratagene) or rTth™ polymerase (Perkin Elmer) with the buffer system provided by the manufacturers, 0.25 μM of each primer, and 0.2 mM of each of the 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min; 57 °C for 1min; and 72 °C for 2 min. The 1.5 kb PCR fragment was digested with Xba I and subcloned into Eco RV-Xba I site of pBluescript. The resulting cDNA clones were fully

sequenced and found to encode two amino acid changes from the published sequences. The first one was a T25N mutation in the N-terminal extracellular domain; the second is an H452Y mutation. Because cDNA clones derived from two independent PCR reactions using Taq polymerase from two different commercial sources (TaqPlus™ from Stratagene and rTth™ Perkin Elmer) contained the same two mutations, these mutations are likely to represent sequence polymorphisms rather than PCR errors. With these exceptions, the nucleic acid (SEQ.ID.NO.: 121) and amino acid (SEQ.ID.NO.: 122) sequences for human $5HT_{2A}$ were thereafter determined and verified.

**38. Serotonin $5HT_{2C}$**

The cDNA encoding endogenous human $5HT_{2C}$ receptor was obtained from human brain poly-$A^+$ RNA by RT-PCR. The 5' and 3' primers were derived from the 5' and 3' untranslated regions and contained the following sequences:

5'-GACCTCGAGGTTGCTTAAGACTGAAGC-3' (SEQ.ID.NO.: 123)

5'-ATTTCTAGACATATGTAGCTTGTACCG-3' (SEQ.ID.NO.: 124)

Nucleic acid (SEQ.ID.NO.: 125) and amino acid (SEQ.ID.NO.: 126) sequences for human $5HT_{2C}$ were thereafter determined and verified.

**39. V28 (GenBank Accession Number: U20350)**

The cDNA for human V28 was generated and cloned into pCMV expression vector as follows: PCR was performed using brain cDNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each of the 4 nucleotides. The cycle condition was 30 cycles of: 94°C for 1 min; 65°C for 1min; and 72°C for 1 min and 20 sec. The 5' PCR primer contained a HindIII site with the sequence:

5'-GGTAAGCTTGGCAGTCCACGCCAGGCCTTC-3' (SEQ.ID.NO.: 127)

and the 3' primer contained an EcoRI site with the sequence:

5'-TCCGAATTCTCTGTAGACACAAGGCTTTGG-3' (SEQ.ID.NO.: 128)

The 1.1 kb PCR fragment was digested with HindIII and EcoRI and cloned into HindIII-EcoRI site of pCMV expression vector. Nucleic acid (SEQ.ID.NO.: 129) and amino acid (SEQ.ID.NO.: 130) sequences for human V28 were thereafter determined and verified.

**Example 2**
**PREPARATION OF NON-ENDOGENOUS HUMAN GPCRS**

**1. Site-Directed Mutagenesis**

Mutagenesis based upon the Human GPCR Proline Marker approach disclosed herein was performed on the foregoing endogenous human GPCRs using Transformer Site-Directed Mutagenesis Kit (Clontech) according to the manufacturer instructions. For this mutagenesis approach, a Mutation Probe and a Selection Marker Probe (unless otherwise indicated, the probe of SEQ.ID.NO.: 132 was the same throughout) were utilized, and the sequences of these for the specified sequences are listed below in Table B (the parenthetical number is the SEQ. ID.NO.). For convenience, the codon mutation incorporated into the human GPCR is also noted, in standard form:

**Table B**

| Receptor Identifier (Codon Mutation) | Mutation Probe Sequence (5'-3') (SEQ.ID.NO.) | Selection Marker Probe Sequence (5'-3') (SEQ.ID.NO.) |
|---|---|---|
| GPR1 (F245K) | GATCTCCAGTAGGCATAAGT GGACAATTCTGG (131) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAG (132) |
| GPR4 (K223A) | AGAAGGCCAAGATCGCGCGG CTGGCCCTCA (133) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| GPR5 (V224K) | CGGCGCCACCGCACGAAAAA GCTCATCTTC | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |

| | | |
|---|---|---|
| | (134) | |
| GPR7 (T250K) | GCCAAGAAGCGGGTGAAGTT CCTGGTGGTGGCA (135) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| GPR8 (T259K) | CAGGCGGAAGGTGAAAGTCC TGGTCCTCGT (136) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| GPR9 (M254K) | CGGCGCCTGCGGGCCAAGCG GCTGGTGGTGGTG (137) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| GPR9-6 (L241K) | CCAAGCACAAAGCCAAGAAA GTGACCATCAC (138) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| GPR10 (F276K) | GCGCCGGCGCACCAAATGCT TGCTGGTGGT (139) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| GPR15 (I240K) | CAAAAAGCTGAAGAAATCTA AGAAGATCATCTTTATTGTCG (140) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| GPR17 (V234K) | CAAGACCAAGGCAAAACGCA TGATCGCCAT (141) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| GPR18 (I231K) | GTCAAGGAGAAGTCCAAAAG GATCATCATC (142) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| GPR20 (M240K) | CGCCGCGTGCGGGCCAAGCA GCTCCTGCTC (143) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| GPR21 (A251K) | CCTGATAAGCGCTATAAAAT GGTCCTGTTTCGA (144) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| GPR22 (F312K) | GAAAGACAAAAGAGAGTCA AGAGGATGTCTTTATTG (145) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| GPR24 (T304K) | CGGAGAAAGAGGGTGAAAC GCACAGCCATCGCC (146) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| GPR30 (L258K) | alternate approach; *see* below | alternate approach; *see* below |
| GPR31 (Q221K) | AAGCTTCAGCGGGCCAAGGC ACTGGTCACC (147) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| GPR32 (K255A) | CATGCCAACCGGCCCGCGAG GCTGCTGCTGGT (279) | ACCAGCAGCAGCCTCGCGGG CCGGTTGGCATG (280) |
| GPR40 (A223K) | CGGAAGCTGCGGGCCAAATG GGTGGCCGGC (265) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| GPR41 | CAGAGGAGGGTGAAGGGGCT GTTGGCG | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |

| | | |
|---|---|---|
| (A223K) | (266) | |
| GPR43 (V221K) | GGCGGCGCCGAGCCAAGGGG CTGGCTGTGG (267) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| APJ (L247K) | alternate approach; *see* below | alternate approach; *see* below |
| BLR1 (V258K) | CAGCGGCAGAAGGCAAAAA GGGTGGCCATC (148) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| CEPR (L258K) | CGGCAGAAGGCGAAGCGCAT GATCCTCGCG (149) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| EBI1 (I262K) | GAGCGCAACAAGGCCAAAA AGGTGATCATC (150) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| EBI2 (L243K) | GGTGTAAACAAAAAGGCTAA AAACACAATTATTCTTATT (151) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| ETBR-LP2 (N358K) | GAGAGCCAGCTCAAGAGCAC CGTGGTG (152) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| GHSR (V262K) | CCACAAGCAAACCAAGAAAA TGCTGGCTGT (153) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| GPCR-CNS (N491K) | CTAGAGAGTCAGATGAAGTG TACAGTAGTGGCAC (155) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| GPR-NGA (I275K) | CGGACAAAAGTGAAAACTAA AAAGATGTTCCTCATT (156) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| H9a and H9b (F236K) | GCTGAGGTTCGCAATAAACT AACCATGTTTGTG (157) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| HB954 (H265K) | GGGAGGCCGAGCTGAAAGCC ACCCTGCTC (158) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| HG38 (V765K) | GGGACTGCTCTATGAAAAAA CACATTGCCCTG (268) | CATCAAGTGTATCATGTGCC AAGTACGCCC (154) |
| HM74 (I230K) | CAAGATCAAGAGAGCCAAAA CCTTCATCATG (159) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| MIG (T273K) | CCGGAGACAAGTGAAGAAG ATGCTGTTTGTC (160) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| OGR1 (Q227K) | GCAAGGACCAGATCAAGCGG CTGGTGCTCA (161) | CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT |
| Serotonin $5HT_{2A}$ (C322K) | alternate approach; *see* below | alternate approach; *see* below |
| Serotonin $5HT_{2C}$ (S310K) | alternate approach; *see* below | alternate approach; *see* below |

| | | |
|---|---|---|
| V28<br>(I230K) | CAAGAAAGCCAAAGCCAAG<br>AAACTGATCCTTCTG<br>(162) | CTCCTTCGGTCCTCCTATCGT<br>TGTCAGAAGT |

The non-endogenous human GPCRs were then sequenced and the derived and verified nucleic acid and amino acid sequences are listed in the accompanying "Sequence Listing" appendix to this patent document, as summarized in Table C below:

**Table C**

| Mutated GPCR | Nucleic Acid Sequence Listing | Amino Acid Sequence Listing |
|---|---|---|
| GPR1<br>(F245K) | SEQ.ID.NO.: 163 | SEQ.ID.NO.: 164 |
| GPR4<br>(K223A) | SEQ.ID.NO.: 165 | SEQ.ID.NO.: 166 |
| GPR5<br>(V224K) | SEQ.ID.NO.: 167 | SEQ.ID.NO.: 168 |
| GPR7<br>(T250K) | SEQ.ID.NO.: 169 | SEQ.ID.NO.: 170 |
| GPR8<br>(T259K) | SEQ.ID.NO.: 171 | SEQ.ID.NO.: 172 |
| GPR9<br>(M254K) | SEQ.ID.NO.: 173 | SEQ.ID.NO.: 174 |
| GPR9-6<br>(L241K) | SEQ.ID.NO.: 175 | SEQ.ID.NO.: 176 |
| GPR10<br>(F276K) | SEQ.ID.NO.: 177 | SEQ.ID.NO.: 178 |
| GPR15<br>(I240K) | SEQ.ID.NO.: 179 | SEQ.ID.NO.: 180 |
| GPR17<br>(V234K) | SEQ.ID.NO.: 181 | SEQ.ID.NO.: 182 |
| GPR18<br>(I231K) | SEQ.ID.NO.: 183 | SEQ.ID.NO.: 184 |
| GPR20<br>(M240K) | SEQ.ID.NO.: 185 | SEQ.ID.NO.: 186 |
| GPR21<br>(A251K) | SEQ.ID.NO.: 187 | SEQ.ID.NO.: 188 |
| GPR22<br>(F312K) | SEQ.ID.NO.: 189 | SEQ.ID.NO.: 190 |
| GPR24<br>(T304K)) | SEQ.ID.NO.: 191 | SEQ.ID.NO.: 192 |
| GPR30 | SEQ.ID.NO.: 193 | SEQ.ID.NO.: 194 |

| | | |
|---|---|---|
| (L258K) | | |
| GPR31 (Q221K) | SEQ.ID.NO.: 195 | SEQ.ID.NO.: 196 |
| GPR32 (K255A) | SEQ.ID.NO.: 269 | SEQ.ID.NO.: 270 |
| GPR40 (A223K) | SEQ.ID.NO.: 271 | SEQ.ID.NO.: 272 |
| GPR41 (A223K) | SEQ.ID.NO.: 273 | SEQ.ID.NO.: 274 |
| GPR43 (V221K) | SEQ.ID.NO.: 275 | SEQ.ID.NO.: 276 |
| APJ (L247K) | SEQ.ID.NO.: 197 | SEQ.ID.NO.: 198 |
| BLR1 (V258K) | SEQ.ID.NO.: 199 | SEQ.ID.NO.: 200 |
| CEPR (L258K) | SEQ.ID.NO.: 201 | SEQ.ID.NO.: 202 |
| EBI1 (I262K) | SEQ.ID.NO.: 203 | SEQ.ID.NO.: 204 |
| EBI2 (L243K) | SEQ.ID.NO.: 205 | SEQ.ID.NO.: 206 |
| ETBR-LP2 (N358K) | SEQ.ID.NO.: 207 | SEQ.ID.NO.: 208 |
| GHSR (V262K) | SEQ.ID.NO.: 209 | SEQ.ID.NO.: 210 |
| GPCR-CNS (N491K) | SEQ.ID.NO.: 211 | SEQ.ID.NO.: 212 |
| GPR-NGA (I275K) | SEQ.ID.NO.: 213 | SEQ.ID.NO.: 214 |
| H9a (F236K) | SEQ.ID.NO.: 215 | SEQ.ID.NO.: 216 |
| H9b (F236K) | SEQ.ID.NO.: 217 | SEQ.ID.NO.: 218 |
| HB954 (H265K) | SEQ.ID.NO.: 219 | SEQ.ID.NO.: 220 |
| HG38 (V765K) | SEQ.ID.NO.: 277 | SEQ.ID.NO.: 278 |
| HM74 (I230K) | SEQ.ID.NO.: 221 | SEQ.ID.NO.: 222 |
| MIG (T273K) | SEQ.ID.NO.: 223 | SEQ.ID.NO.: 224 |
| OGR1 (Q227K) | SEQ.ID.NO.: 225 | SEQ.ID.NO.: 226 |
| Serotonin $5HT_{2A}$ (C322K) | SEQ.ID.NO.: 227 | SEQ.ID.NO.: 228 |
| Serotonin $5HT_{2C}$ (S310K) | SEQ.ID.NO.: 229 | SEQ.ID.NO.: 230 |
| V28 (I230K) | SEQ.ID.NO.: 231 | SEQ.ID.NO.: 232 |

**2. Alternate Mutation Approaches for Employment of the Proline Marker Algorithm: APJ; Serotonin $5HT_{2A}$; Serotonin $5HT_{2C}$; and GPR30**

Although the above site-directed mutagenesis approach is particularly preferred, other approaches can be utilized to create such mutations; those skilled in the art are readily credited with selecting approaches to mutating a GPCR that fits within the particular needs of the artisan.

***a. APJ***

Preparation of the non-endogenous, human APJ receptor was accomplished by mutating L247K. Two oligonucleotides containing this mutation were synthesized:

5'- GGCTTAAGAGCATCATCGTGGTGCTGGTG-3' (SEQ.ID.NO.: 233 )

5'-GTCACCACCAGCACCACGATGATGCTCTTAAGCC-3' (SEQ.ID.NO.: 234)

The two oligonucleotides were annealed and used to replace the NaeI-BstEII fragment of human, endogenous APJ to generate the non-endogenous, version of human APJ.

***b. Serotonin $5HT_{2A}$***

cDNA containing the point mutation C322K was constructed by utilizing the restriction enzyme site Sph I which encompasses amino acid 322. A primer containing the C322K mutation:

5'-CAAAGAAAGTACTGGGCATCGTCTTCTTCCT-3' (SEQ.ID.NO: 235)

was used along with the primer from the 3' untranslated region of the receptor:

5'-TGCTCTAGATTCCAGATAGGTGAAAA CTTG-3' (SEQ.ID.NO.: 236)

to perform PCR (under the conditions described above). The resulting PCR fragment was then used to replace the 3' end of endogenous $5HT_{2A}$ cDNA through the T4 polymerase blunted Sph I site.

***c. Serotonin $5HT_{2C}$***

The cDNA containing a S310K mutation was constructed by replacing the Sty I restriction fragment containing amino acid 310 with synthetic double stranded oligonucleotides that encode

the desired mutation. The sense strand sequence utilized had the following sequence:

5'-CTAGGGGCACCATGCAGGCTATCAACAATGAAAGAAAAGCTAAGAAAGTC-3'

(SEQ. ID.NO.: 237)

and the antisense strand sequence utilized had the following sequence:

5'-CAAGGACTTTCTTAGCTTTTCTTTCATTGTTGATAGCCTGCATGGTGCCC-3' (SEQ. ID. NO.: 238)

*d. GPR30*

Prior to generating non-endogenous GPR30, several independent pCR2.1/GPR30 isolates were sequenced in their entirety in order to identify clones with no PCR-generated mutations. A clone having no mutations was digested with EcoR1 and the endogenous GPR30 cDNA fragment was transferred into the CMV-driven expression plasmid pCI-neo (Promega), by digesting pCI-Neo with EcoRI and subcloning the EcoRI-liberated GPR30 fragment from pCR2.1/GPR30, to generate pCI/GPR30. Thereafter, the leucine at codon 258 was mutated to a lysine using a Quick-Change™ Site-Directed Mutagenesis Kit (Stratagene, #200518), according to manufacturer's instructions, and the following primers:

5'-CGGCGGCAGAAGGCGAAACGCATGATCCTCGCGGT-3' (SEQ.ID.NO.: 239) and

5'-ACCGCGAGGATCATGCGTTTCGCCTTCTGC CGCCG-3' (SEQ.ID.NO.: 240)

**Example 3**

**Receptor (Endogenous and Mutated) Expression**

Although a variety of cells are available to the art for the expression of proteins, it is most preferred that mammalian cells be utilized. The primary reason for this is predicated upon practicalities, *i.e.*, utilization of, *e.g.*, yeast cells for the expression of a GPCR, while possible,

introduces into the protocol a non-mammalian cell which may not (indeed, in the case of yeast, does not) include the receptor-coupling, genetic-mechanism and secretary pathways that have evolved for mammalian systems – thus, results obtained in non-mammalian cells, while of potential use, are not as preferred as that obtained from mammalian cells. Of the mammalian cells, COS-7, 293 and 293T cells are particularly preferred, although the specific mammalian cell utilized can be predicated upon the particular needs of the artisan.

Unless otherwise noted herein, the following protocol was utilized for the expression of the endogenous and non-endogenous human GPCRs. Table D lists the mammalian cell and number utilized (per 150mm plate) for GPCR expression.

**Table D**

| Receptor Name (Endogenous or Non-Endogenous) | Mammalian Cell (Number Utilized) |
|---|---|
| GPR17 | 293 ($2 \times 10^4$) |
| GPR30 | 293 ($4 \times 10^4$) |
| APJ | COS-7 ($5X10^6$) |
| ETBR-LP2 | 293 ($1 \times 10^7$)<br>293T ($1 \times 10^7$) |
| GHSR | 293 ($1 \times 10^7$)<br>293T ($1 \times 10^7$) |
| MIG | 293 ($1 \times 10^7$) |
| Serotonin $5HT_{2A}$ | 293T ($1 \times 10^7$) |
| Serotonin $5HT_{2c}$ | 293T ($1 \times 10^7$) |

On day one, mammalian cells were plated out. On day two, two reaction tubes were prepared (the proportions to follow for each tube are per plate): tube A was prepared by mixing 20μg DNA (*e.g.*, pCMV vector; pCMV vector with endogenous receptor cDNA, and pCMV vector with non-endogenous receptor cDNA.) in 1.2ml serum free DMEM (Irvine Scientific,

Irvine, CA); tube B was prepared by mixing 120µl lipofectamine (Gibco BRL) in 1.2ml serum free DMEM. Tubes A and B were then admixed by inversions (several times), followed by incubation at room temperature for 30-45min. The admixture is referred to as the "transfection mixture". Plated cells were washed with 1XPBS, followed by addition of 10ml serum free DMEM. 2.4ml of the transfection mixture was then added to the cells, followed by incubation for 4hrs at 37°C/5% $CO_2$. The transfection mixture was then removed by aspiration, followed by the addition of 25ml of DMEM/10% Fetal Bovine Serum. Cells were then incubated at 37°C/5% $CO_2$. After 72hr incubation, cells were then harvested and utilized for analysis.

1. **Gi-Coupled Receptors: Co-Transfection with Gs-Coupled Receptors**

In the case of GPR30, it has been determined that this receptor couples the G protein Gi. Gi is known to inhibit the enzyme adenylyl cyclase, which is necessary for catalyzing the conversion of ATP to cAMP. Thus, a non-endogenous, constitutively activated form of GPR30 would be expected to be associated with decreased levels of cAMP. Assay confirmation of a non-endogenous, constitutively activated form of GPR30 directly via measurement of decreasing levels of cAMP, while viable, can be preferably measured by cooperative use of a Gs-coupled receptor. For example, a receptor that is Gs-coupled will stimulate adenylyl cyclase, and thus will be associated with an increase in cAMP. The assignee of the present application has discovered that the orphan receptor GPR6 is an endogenous, constitutively activated GPCR. GPR6 couples to the Gs protein. Thus when co-transfected, one can readily verify that a putative GPR30-mutation leads to constitutive activation thereof: *i.e.*, an endogenous, constitutively activated GPR6/endogenous, non-constitutively activated GPR30 cell will evidence an elevated level of cAMP when compared with an endogenous, constitutively active GPR6/non-endogenous, constitutively activated GPR30 (the latter evidencing a comparatively lower level of cAMP).

Assays that detect cAMP can be utilized to determine if a candidate compound is *e.g.*, an inverse agonist to a Gs-associated receptor (*i.e.*, such a compound would decrease the levels of cAMP) or a Gi-associated receptor (or a Go-associated receptor) (*i.e.*, such a candidate compound would increase the levels of cAMP). A variety of approaches known in the art for measuring cAMP can be utilized; a preferred approach relies upon the use of anti-cAMP antibodies. Another approach, and most preferred, utilizes a whole cell second messenger reporter system assay. Promoters on genes drive the expression of the proteins that a particular gene encodes. Cyclic AMP drives gene expression by promoting the binding of a cAMP-responsive DNA binding protein or transcription factor (CREB) which then binds to the promoter at specific sites called cAMP response elements and drives the expression of the gene. Reporter systems can be constructed which have a promoter containing multiple cAMP response elements before the reporter gene, *e.g.*, β-galactosidase or luciferase. Thus, an activated receptor such as GPR6 causes the accumulation of cAMP which then activates the gene and expression of the reporter protein. Most preferably, 293 cells are co-transfected with GPR6 (or another Gs-linked receptor) and GPR30 (or another Gi-linked receptor) plasmids, preferably in a 1:1 ratio, most preferably in a 1:4 ratio. Because GPR6 is an endogenous, constitutively active receptor that stimulates the production of cAMP, GPR6 strongly activates the reporter gene and its expression. The reporter protein such as β-galactosidase or luciferase can then be detected using standard biochemical assays (Chen et al. 1995). Co-transfection of endogenous, constitutively active GPR6 with endogenous, non-constitutively active GPR30 evidences an increase in the luciferase reporter protein. Conversely, co-transfection of endogenous, constitutively active GPR6 with non-endogenous, constitutively active GPR30 evidences a drastic decrease in expression of luciferase. Several reporter plasmids are known and available in the art for measuring a second messenger assay. It is considered well within the

skilled artisan to determine an appropriate reporter plasmid for a particular gene expression based primarily upon the particular need of the artisan. Although a variety of cells are available for expression, mammalian cells are most preferred, and of these types, 293 cells are most preferred. 293 cells were transfected with the reporter plasmid pCRE-Luc/GPR6 and non-endogenous, constitutively activated GPR30 using a Mammalian Transfection™ Kit (Stratagene, #200285) $CaPO_4$ precipitation protocol according to the manufacturer's instructions (*see*, 28 Genomics 347 (1995) for the published endogenous GPR6 sequence). The precipitate contained 400ng reporter, 80ng CMV-expression plasmid (having a 1:4 GPR6 to endogenous GPR30 or non-endogenous GPR30 ratio) and 20ng CMV-SEAP (a transfection control plasmid encoding secreted alkaline phosphatase). 50% of the precipitate was split into 3 wells of a 96-well tissue culture dish (containing $4X10^4$ cells/well); the remaining 50% was discarded. The following morning, the media was changed. 48 hr after the start of the transfection, cells were lysed and examined for luciferase activity using a Luclite™ Kit (Packard, Cat. # 6016911) and Trilux 1450 Microbeta™ liquid scintillation and luminescence counter (Wallac) as per the vendor's instructions. The data were analyzed using GraphPad Prism 2.0a (GraphPad Software Inc.).

With respect to GPR17, which has also been determined to be Gi-linked, a modification of the foregoing approach was utilized, based upon, *inter alia*, use of another Gs-linked endogenous receptor, GPR3 (*see* 23 Genomics 609 (1994) and 24 Genomics 391 (1994)). Most preferably, 293 cells are utilized. These cells were plated-out on 96 well plates at a density of 2 x $10^4$ cells per well and were transfected using Lipofectamine Reagent (BRL) the following day according to manufacturer instructions. A DNA/lipid mixture was prepared for each 6-well transfection as follows: 260ng of plasmid DNA in 100µl of DMEM were gently mixed with 2µl of lipid in 100µl of DMEM (the 260ng of plasmid DNA consisted of 200ng of a 8xCRE-Luc

reporter plasmid (*see* below), 50ng of pCMV comprising endogenous receptor or non-endogenous receptor or pCMV alone, and 10ng of a GPRS expression plasmid (GPRS in pcDNA3 (Invitrogen)). The 8XCRE-Luc reporter plasmid was prepared as follows: vector SRIF-β-gal was obtained by cloning the rat somatostatin promoter (-71/+51) at BglV-HindIII site in the pβgal-Basic Vector (Clontech). Eight (8) copies of cAMP response element were obtained by PCR from an adenovirus template AdpCF126CCRE8 (*see* 7 Human Gene Therapy 1883 (1996)) and cloned into the SRIF-β-gal vector at the Kpn-BglV site, resulting in the 8xCRE-β-gal reporter vector. The 8xCRE-Luc reporter plasmid was generated by replacing the beta-galactosidase gene in the 8xCRE-β-gal reporter vector with the luciferase gene obtained from the pGL3-basic vector (Promega) at the HindIII-BamHI site. Following 30min. incubation at room temperature, the DNA/lipid mixture was diluted with 400 µl of DMEM and 100µl of the diluted mixture was added to each well. 100 µl of DMEM with 10% FCS were added to each well after a 4hr incubation in a cell culture incubator. The next morning the transfected cells were changed with 200 µl/well of DMEM with 10% FCS. Eight (8) hours later, the wells were changed to 100 µl /well of DMEM without phenol red, after one wash with PBS. Luciferase activity were measured the next day using the LucLite™ reporter gene assay kit (Packard) following manufacturer instructions and read on a 1450 MicroBeta™ scintillation and luminescence counter (Wallac).

Figure 4 evidences that constitutively active GPR30 inhibits GPR6-mediated activation of CRE-Luc reporter in 293 cells. Luciferase was measured at about 4.1 relative light units in the expression vector pCMV. Endogenous GPR30 expressed luciferase at about 8.5 relative light units, whereas the non-endogenous, constitutively active GPR30 (L258K), expressed luciferase at about 3.8 and 3.1 relative light units, respectively. Co-transfection of endogenous GPR6 with endogenous GPR30, at a 1:4 ratio, drastically increased luciferase

expression to about 104.1 relative light units. Co-transfection of endogenous GPR6 with non-endogenous GPR30 (L258K), at the same ratio, drastically decreased the expression, which is evident at about 18.2 and 29.5 relative light units, respectively. Similar results were observed with respect to GPR17 with respect to co-transfection with GPR3, as set forth in Figure 5.

**Example 3**
**ASSAYS FOR DETERMINATION OF CONSTITUTIVE ACTIVITY OF NON-ENDOGENOUS GPCRS**

**A. Membrane Binding Assays**

**1. [$^{35}$S]GTPγS Assay**

When a G protein-coupled receptor is in its active state, either as a result of ligand binding or constitutive activation, the receptor couples to a G protein and stimulates the release of GDP and subsequent binding of GTP to the G protein. The alpha subunit of the G protein-receptor complex acts as a GTPase and slowly hydrolyzes the GTP to GDP, at which point the receptor normally is deactivated. Constitutively activated receptors continue to exchange GDP for GTP. The non-hydrolyzable GTP analog, [$^{35}$S]GTPγS, can be utilized to demonstrate enhanced binding of [$^{35}$S]GTPγS to membranes expressing constitutively activated receptors. The advantage of using [$^{35}$S]GTPγS binding to measure constitutive activation is that: (a) it is generically applicable to all G protein-coupled receptors; (b) it is proximal at the membrane surface making it less likely to pick-up molecules which affect the intracellular cascade.

The assay utilizes the ability of G protein coupled receptors to stimulate [$^{35}$S]GTPγS binding to membranes expressing the relevant receptors. The assay can, therefore, be used in the direct identification method to screen candidate compounds to known, orphan and constitutively activated G protein-coupled receptors. The assay is generic and has application

to drug discovery at all G protein-coupled receptors.

The [$^{35}$S]GTPγS assay was incubated in 20 mM HEPES and between 1 and about 20mM $MgCl_2$ (this amount can be adjusted for optimization of results, although 20mM is preferred) pH 7.4, binding buffer with between about 0.3 and about 1.2 nM [$^{35}$S]GTPγS (this amount can be adjusted for optimization of results, although 1.2 is preferred ) and 12.5 to 75 μg membrane protein (*e.g.* COS-7 cells expressing the receptor; this amount can be adjusted for optimization, although 75μg is preferred) and 1 μM GDP (this amount can be changed for optimization) for 1 hour. Wheatgerm agglutinin beads (25 μl; Amersham) were then added and the mixture was incubated for another 30 minutes at room temperature. The tubes were then centrifuged at 1500 x g for 5 minutes at room temperature and then counted in a scintillation counter.

A less costly but equally applicable alternative has been identified which also meets the needs of large scale screening. Flash plates™ and Wallac™ scintistrips may be utilized to format a high throughput [$^{35}$S]GTPγS binding assay. Furthermore, using this technique, the assay can be utilized for known GPCRs to simultaneously monitor tritiated ligand binding to the receptor at the same time as monitoring the efficacy via [$^{35}$S]GTPγS binding. This is possible because the Wallac beta counter can switch energy windows to look at both tritium and $^{35}$S-labeled probes. This assay may also be used to detect other types of membrane activation events resulting in receptor activation. For example, the assay may be used to monitor $^{32}$P phosphorylation of a variety of receptors (both G protein coupled and tyrosine kinase receptors). When the membranes are centrifuged to the bottom of the well, the bound [$^{35}$S]GTPγS or the $^{32}$P-phosphorylated receptor will activate the scintillant which is coated of the wells. Scinti® strips (Wallac) have been used to demonstrate this principle. In addition, the assay also has utility for measuring ligand binding to receptors using radioactively labeled ligands. In a similar manner, when the radiolabeled bound

ligand is centrifuged to the bottom of the well, the scintistrip label comes into proximity with the radiolabeled ligand resulting in activation and detection.

Representative results of graph comparing Control (pCMV), Endogenous APJ and Non-Endogenous APJ, based upon the foregoing protocol, are set forth in Figure 6.

**2. Adenylyl Cyclase**

A Flash Plate™ Adenylyl Cyclase kit (New England Nuclear; Cat. No. SMP004A) designed for cell-based assays was modified for use with crude plasma membranes. The Flash Plate wells contain a scintillant coating which also contains a specific antibody recognizing cAMP. The cAMP generated in the wells was quantitated by a direct competition for binding of radioactive cAMP tracer to the cAMP antibody. The following serves as a brief protocol for the measurement of changes in cAMP levels in membranes that express the receptors.

Transfected cells were harvested approximately three days after transfection. Membranes were prepared by homogenization of suspended cells in buffer containing 20mM HEPES, pH 7.4 and 10mM $MgCl_2$. Homogenization was performed on ice using a Brinkman Polytron™ for approximately 10 seconds. The resulting homogenate was centrifuged at 49,000 X g for 15 minutes at 4°C. The resulting pellet was then resuspended in buffer containing 20mM HEPES, pH 7.4 and 0.1 mM EDTA, homogenized for 10 seconds, followed by centrifugation at 49,000 X g for 15 minutes at 4°C. The resulting pellet can be stored at -80°C until utilized. On the day of measurement, the membrane pellet was slowly thawed at room temperature, resuspended in buffer containing 20mM HEPES, pH 7.4 and 10mM $MgCL_2$ (these amounts can be optimized, although the values listed herein are prefereed), to yield a final protein concentration of 0.60mg/ml (the resuspended membranes were placed on ice until use).

cAMP standards and Detection Buffer (comprising 2 μCi of tracer [$^{125}$I cAMP (100 μl] to

11 ml Detection Buffer) were prepared and maintained in accordance with the manufacturer's instructions. Assay Buffer was prepared fresh for screening and contained 20mM HEPES, pH 7.4, 10mM $MgCl_2$, 20mM (Sigma), 0.1 units/ml creatine phosphokinase (Sigma), 50 μM GTP (Sigma), and 0.2 mM ATP (Sigma); Assay Buffer can be stored on ice until utilized. The assay was initiated by addition of 50ul of assay buffer followed by addition of 50ul of membrane suspension to the NEN Flash Plate. The resultant assay mixture is incubated for 60 minutes at room temperature followed by addition of 100ul of detection buffer. Plates are then incubated an additional 2-4 hours followed by counting in a Wallac MicroBeta scintillation counter. Values of cAMP/well are extrapolated from a standard cAMP curve which is contained within each assay plate. The foregoing assay was utilized with respect to analysis of MIG.

**B. Reporter-Based Assays**

**1. CREB Reporter Assay (Gs-associated receptors)**

A method to detect Gs stimulation depends on the known property of the transcription factor CREB, which is activated in a cAMP-dependent manner. A PathDetect CREB trans-Reporting System (Stratagene, Catalogue # 219010) was utilized to assay for Gs coupled activity in 293 or 293T cells. Cells were transfected with the plasmids components of this above system and the indicated expression plasmid encoding endogenous or mutant receptor using a Mammalian Transfection Kit (Stratagene, Catalogue #200285) according to the manufacurer's instructions. Briefly, 400 ng pFR-Luc (luciferase reporter plasmid containing Gal4 recognition sequences), 40 ng pFA2-CREB (Gal4-CREB fusion protein containing the Gal4 DNA-binding domain), 80 ng CMV-receptor expression plasmid (comprising the receptor) and 20 ng CMV-SEAP (secreted alkaline phosphatase expression plasmid; alkaline phosphatase activity is measured in the media of transfected cells to control for variations in

transfection efficiency between samples) were combined in a calcium phosphate precipitate as per the Kit's instructions. Half of the precipitate was equally distributed over 3 wells in a 96-well plate, kept on the cells overnight, and replaced with fresh medium the following morning. Forty-eight (48) hr after the start of the transfection, cells were treated and assayed for luciferase activity as set forth with resepct to the GPR30 system, above. This assay was used with respect to GHSR.

**2. AP1 reporter assay (Gq-associated receptors)**

Ae method to detect Gq stimulation depends on the known property of Gq-dependent phospholipase C to cause the activation of genes containing AP1 elements in their promoter. A Pathdetect AP-1 cis-Reporting System (Stratagene, Catalogue # 219073) was utilized following the protocl set forth above with respect to the CREB reporter assay, except that the components of the calcium phosphate precipitate were 410 ng pAP1-Luc, 80 ng receptor expression plasmid, and 20 ng CMV-SEAP. This assay was used with respect to ETBR-LP2

**C. Intracellular IP3 Accumulation Assay**

On day 1, cells comprising the serotonin receptors (endogenous and mutated) were plated onto 24 well plates, usually $1x10^5$ cells/well. On day 2 cells were transfected by firstly mixing 0.25ug DNA in 50 ul serumfree DMEM/well and 2 ul lipofectamine in 50 µl serumfree DMEM/well. The solutions were gently mixed and incubated for 15-30 min at room temperature. Cells were washed with 0.5 ml PBS and 400 µl of serum free media was mixed with the transfection media and added to the cells. The cells were then incubated for 3-4 hrs at 37°C/5%$CO_2$ and then the transfection media was removed and replaced with 1ml/well of regular growth media. On day 3 the cells were labeled with $^3$H-myo-inositol. Briefly, the media was removed the cells were washed with 0.5 ml PBS. Then 0.5 ml inositol-free/serumfree media ( GIBCO BRL) was added/well with 0.25 µCi of $^3$H-myo-inositol / well

and the cells were incubated for 16-18 hrs o/n at 37°C/5%$CO_2$ . On Day 4 the cells were washed with 0.5 ml PBS and 0.45 ml of assay medium was added containing inositol-free/serum free media 10 μM pargyline 10 mM lithium chloride or 0.4 ml of assay medium and 50 ul of 10x ketanserin (ket) to final concentration of 10μM. The cells were then incubated for 30 min at 37°C. The cells were then washed with 0.5 ml PBSand 200 ul of fresh/icecold stop solution (1M KOH; 18 mM Na-borate; 3.8 mM EDTA) was added/well. The solution was kept on ice for 5-10 min or until cells were lysed and then neutralized by 200 μl of fresh/ice cold neutralization sol. (7.5 % HCL). The lysate was then transferred into 1.5 ml eppendorf tubes and 1 ml of chloroform/methanol (1:2) was added/tube. The solution was vortexed for 15 sec and the upper phase was applied to a Biorad AG1-X8 anion exchange resin (100-200 mesh). Firstly, the resin was washed with water at 1:1.25 W/V and 0.9 ml of upper phase was loaded onto the column. The column was washed with 10 mls of 5 mM myo-inositol and 10 ml of 5 mM Na-borate/60mM Na-formate. The inositol tris phosphates were eluted into scintillation vials containing 10 ml of scintillation cocktail with 2 ml of 0.1 M formic acid/ 1 M ammonium formate. The columns were regenerated by washing with 10 ml of 0.1 M formic acid/3M ammonium formate and rinsed twice with dd $H_2O$ and stored at 4°C in water.

Figure 7 provides an illustration of IP3 production from the human 5-$HT_{2A}$ receptor that incorporates the C322K mutation. While these results evidence that the Proline Mutation Algorithm approach constitutively activates this receptor, for purposes of using such a receptor for screening for identification of potential therapeutics, a more robust difference would be preferred. However, because the activated receptor can be utilized for understanding and elucidating the role of constitutive activation and for the identification of compounds that

can be further examined, we believe that this difference is itself useful in differentiating between the endogenous and non-endogenous versions of the human $5HT_{2A}$ receptor.

**D. Result Summary**

The results for the GPCRs tested are set forth in Table E where the Per-Cent Increase indicates the percentage difference in results observed for the non-endogenous GPCR as compared to the endogenous GPCR; these values are followed by parenthetical indications as to the type of assay utilized. Additionally, the assay sytem utilized is parenthetically listed (and, in cases where different Host Cells were used, both are listed). As these results indicate, a variety of assays can be utilized to determine constitutive activity of the non-endogenous versions of the human GPCRs. Those skilled in the art, based upon the foregoing and with reference to information available to the art, are creditied with theability to selelect and/ot maximize a particular assay approach that suites the particualr needs of theinvestigator.

**Table E**

| Receptor Identifier (Codon Mutation) | Per-Cent Difference |
|---|---|
| GPR17 (V234K) | 74.5 (CRE-Luc) |
| GPR30 (L258K) | 71.6 (CREB) |
| APJ (L247K) | 49.0 (GTPγS) |
| ETBR-LP2 (N358K) | 48.4(AP1-Luc - 293)<br>61.1(AP1-Luc – 293T) |
| GHSR (V262K) | 58.9(CREB – 293)<br>35.6(CREB – 293T) |

| | |
|---|---|
| MIG (I230K) | 39 (cAMP) |
| Serotonin $5HT_{2A}$ (C322K) | 33.2 ($IP_3$) |
| Serotonin $5HT_{2C}$ (S310K) | 39.1($IP_3$) |

**Example 6**

**Tissue Distribution of Endogenous Orphan GPCRs**

Using a commercially available human-tissue dot-blot format, endogenous orphan GPCRs were probed for a determination of the areas where such receptors are localized. Except as indicate below, the entire receptor cDNA (radiolabelled) was used as the probe: radiolabeled probe was generated using the complete receptor cDNA (excised from the vector) using a Prime-It II™ Random Primer Labeling Kit (Stratagene, #300385), according to manufacturer's instructions. A human RNA Master Blot™ (Clontech, #7770-1) was hybridized with the GPCR radiolabeled probe and washed under stringent conditions according manufacturer's instructions. The blot was exposed to Kodak BioMax Autoradiography film overnight at -80°C.

Representative dot-blot format results are presented in Figure 8 for GPR1 (8A), GPR30 (8B), and APJ (8C), with results being summarized for all receptors in Table F

**Table F**

| GPCR | Tissue Distribution (highest levels, relative to other tissues in the dot-blot) |
|---|---|
| GPR1 | Placenta, Ovary, Adrenal |

| | |
|---|---|
| GPR4 | Broad; highest in Heart, Lung, Adrenal, Thyroid, Spinal Cord |
| GPR5 | Placenta, Thymus, Fetal Thymus<br>Lesser levels in spleen, fetal spleen |
| GPR7 | Liver, Spleen, Spinal Cord, Placenta |
| GPR8 | No expression detected |
| GPR9-6 | Thymus, Fetal Thymus<br>Lesser levels in Small Intestine |
| GPR18 | Spleen, Lymph Node, Fetal Spleen, Testis |
| GPR20 | Broad |
| GPR21 | Broad; very low abundance |
| GPR22 | Heart, Fetal Heart<br>Lesser levels in Brain |
| GPR30 | Stomach |
| GPR31 | Broad |
| BLR1 | Spleen |
| CEPR | Stomach, Liver, Thyroid, Putamen |
| EBI1 | Pancreas<br>Lesser levels in Lymphoid Tissues |
| EBI2 | Lymphoid Tissues, Aorta, Lung, Spinal Cord |
| ETBR-LP2 | Broad; Brain Tissue |
| GPCR-CNS | Brain<br>Lesser levels in Testis, Placenta |
| GPR-NGA | Pituitary<br>Lesser levels in Brain |
| H9 | Pituitary |
| HB954 | Aorta, Cerebellum<br>Lesser levels in most other tissues |
| HM74 | Spleen, Leukocytes, Bone marrow, Mammary Glands, Lung, Trachea |
| MIG | Low levels in Kidney, Liver, Pancreas, Lung, Spleen |
| ORG1 | Pituitary, Stomach, Placenta |
| V28 | Brain, Spleen, Peripheral Leukocytes |

Based upon the foregoing information, it is noted that human GPCRs can also be assessed for distribution in diseased tissue; comparative assessments between "normal" and diseased tissue can then be utilized to determine the potential for over-expression or under-expression of a particular receptor in a diseased state. In those circumstances where it is desirable to utilize the non-endogenous versions of the human GPCRs for the purpose of screening to directly identify

candidate compounds of potential therapeutic relevance, it is noted that inverse agonists are useful in the treatment of diseases and disorders where a particular human GPCR is over-expressed, whereas agonists or partial agonists are useful in the treatment of diseases and disorders where a particular human GPCR is under-expressed.

As desired, more detailed, cellular localization of the recepotrs, using techniques well-known to those in the art (*e.g.*, in-situ hybridization) can be utilized to identify particualr cells within these tissues where the receptor of interest is expressed.

It is intended that each of the patents, applications, and printed publications mentioned in this patent document be hereby incorporated by reference in their entirety.

As those skilled in the art will appreciate, numerous changes and modifications may be made to the preferred embodiments of the invention without departing from the spirit of the invention. It is intended that all such variations fall within the scope of the invention.

Although a variety of expression vectors are available to those in the art, for purposes of utilization for both the endogenous and non-endogenous human GPCRs, it is most preferred that the vector utilized be pCMV. This vector has been deposited with the American Type Culture Collection (ATCC) on October 13, 1998 (10801 University Blvd., Manassas, VA 20110-2209 USA) under the provsions of the Budapest Treaty for the International Recognition of the Deposit of Microorganisms for the Purpose of patent Procedure. The vector was tested by the ATCC on _____ __, 1998 and determined to be viable on _____ __, 1998. The ATCC has assigned the following deposit number to pCMV: _____.

## CLAIMS

What is claimed is:

1. A constitutively active, non-endogenous version of an endogenous human orphan G protein-coupled receptor (GPCR) comprising the following amino acid residues (carboxy-terminus to amino-terminus orientation) transversing the transmembrane-6 (TM6) and intracellular loop-3 (IC3) regions of the non-endogenous GPCR:

$$P^1\ AA_{15}\ X$$

wherein:

(1) $P^1$ is an amino acid residue located within the TM6 region of the non-endogenous GPCR, where $P^1$ is selected from the group consisting of (i) the endogenous orphan GPCR proline residue, and (ii) a non-endogenous amino acid residue other than proline;

(2) $AA_{15}$ are 15 amino acid residues selected from the group consisting of (a) the 15 endogenous amino acid residues of the endogenous orphan GPCR, (b) 15 non-endogenous amino acid residues, and (c) a combination of 15 amino acid residues, the combination comprising at least one endogenous amino acid residue of the endogenous orphan GPCR and at least one non-endogenous amino acid residue, excepting that none of the 15 endogenous amino acid residues that are positioned within the TM6 region of the GPCR is proline; and

(2) X is a non-endogenous amino acid residue located within the IC3 region of said non-endogenous GPCR.

2. The non-endogenous human GPCR of claim 1 wherein $P^1$ is the endogenous proline

residue.

3. The non-endogenous human GPCR of claim1 wherein $P^1$ is a non-endogenous amino acid residue other than a proline residue.

4. The non-endogenous human GPCR of claim 1 wherein $AA_{15}$ are the 15 endogenous amino acid residues of the endogenous GPCR.

5. The non-endogenous human GPCR of claim 1 wherein X is selected from the group consisting of lysine, hisitidine, arganine and alanine residues, excepting that when the endogenous amino acid in position X of said endogenous human GPCR is lysine, X is selected from the group consisting of histidine, arginine and alanine.

6. The non-endogenous human GPCR of claim 1 wherein X is a lysine residue, excepting that when the endogenous amino acid in position X of said endogenous human GPCR is lysine, X is an amino acid other than lysine.

7. The non-endogenous human GPCR of claim 4 wherein X is a lysine residue, excepting that when the endogenous amino acid in position X of said endogenous human GPCR is lysine, X is an amino acid other than lysine.

8. The non-endogenous, human GPCR of claim 1 wherein $P^1$ is a proline residue and X is a lysine residue, excepting that when the endogenous amino acid in position X of said endogenous human GPCR is lysine, X is an amino acid other than lysine.

9. A host cell comprising the non-endogenous human GPCR of claim 1.

10. The material of claim 9 wherein said host cell is of mammalian origin.

11. The non-endogenous human GPCR of claim 1 in a purified and isolated form.

12. A nucleic acid sequence encoding a constitutively active, non-endogenous version of an endogenous human orphan G protein-coupled receptor (GPCR) comprising the following

nucleic acid sequence region transversing the transmembrane-6 (TM6) and intracellular loop-3 (IC3) regions of the orphan GPCR:

$$3'\text{-}P^{codon}\ (AA\text{-}codon)_{15}\ X_{codon}\text{-}5'$$

wherein:

(1) $P^{codon}$ is a nucleic acid encoding region within the TM6 region of the non-endogenous GPCR, where $P^{codon}$ encodes an amino acid selected from the group consisting of (i) the endogenous GPCR proline residue, and (ii) a non-endogenous amino acid residue other than proline;

(2) $(AA\text{-}codon)_{15}$ are 15 codons encoding 15 amino acid residues selected from the group consisting of (a) the 15 endogenous amino acid residues of the endogenous orphan GPCR, (b) 15 non-endogenous amino acid residues, and (c) a combination of 15 amino acid residues, the combination comprising at least one endogenous amino acid residue of the endogenous orphan GPCR and at least one non-endogenous amino acid residue, excepting that none of the 15 endogenous amino acid residues that are positioned within the TM6 region of the orphan GPCR is proline; and

(3) $X_{codon}$ is a nucleic acid encoding region residue located within the IC3 region of said non-endogenous human GPCR, where $X_{codon}$ encodes a non-endogenous amino acid.

13. The nucleic acid sequence of claim 12 wherein $P^{codon}$ encodes an endogenous proline residue.

14. The nucleic acid sequence of claim 12 wherein $P^{codon}$ encodes a non-endogenous

amino acid residue other than a proline residue.

15. The nucleic acid sequence of claim 12 wherein $X_{codon}$ encodes a non-endogenous amino acid selected from the group consisting of lysine, histidine, arginine and alanine, excepting that when the endogenous amino acid in position X of said endogenous human GPCR is lysine, $X_{codon}$ encodes an amino acid selected from the group consisiting of histidine, arginine and alanine.

16. The nucleic acid sequence of claim 13 wherein $X_{codon}$ encodes a non-endogenous lysine amino acid excepting that when the endogenous amino acid in position X of said endogenous human GPCR is lysine, $X_{codon}$ encodes an amino acid selected from the group consisiting of histidine, arginine and alanine.

17. The nucleic acid sequence of claim 12 wherein $X_{codon}$ is selected from the group consisting of AAA, AAG, GCA, GCG, GCC and GCU.

18. The nucleic acid sequence of claim 12 wherein $X_{codon}$ is selected from the group consisting of AAA and AAG.

19. The nucleic acid sequence of claim 12 wherein $P^{codon}$ is selected from the group consisting of CCA, CCC, CCG and CCU, and $X_{codon}$ is selected from the group consisting of AAA and AAG.

20. A vector comprising the nucleic acid sequence of claim 12.

21. A plasmid comprising the nucleic acid sequence of claim 12.

22. A host cell comprising the nucleic acid sequence of claim 21.

23. The nucleic acid sequence of claim 12 in a purified and isolated form.

24. A method for selecting for alteration an endogenous amino acid residue within the third intracellular loop of a human G protein-coupled receptor ("GPCR"), said receptor

comprising a transmembrane 6 region and an intracellular loop 3 region, which endogenous amino acid, when altered to a non-endogenous amino acid, constitutively activates said human GPCR, comprising the following steps:

(a) identifying an endogenous proline residue within the transmembrane 6 region of a human GPCR;

(b) identifying, by moving in a direction of the carboxy-terminus region of said GPCR towards the amino-terminus region of said GPCR, the endogenous, 16$^{th}$ amino acid residue from said proline residue;

(c) altering the endogenous residue of step (b) to a non-endogenous amino acid residue to create a non-endogenous version of an endogenous human GPCR; and

(d) determining whether the non-endogenous human GPCR of step (c) is constitutively active.

25. The method of claim 24 wherein the amino acid residue that is two residues from said proline residue in the transmembrane 6 region, in a carboxy-terminus to amino-terminus direction, is tryptophan.

26. A constitutively active, non-endogenous human GPCR produced by the process of claim 24.

27. A constitutively active, non-endogenous human GPCR produced by the process of claim 25.

28. An algorithmic approach for creating a non-endogenous, constitutively active version of an endogenous human G protein coupled receptor (GPCR), said endogenous GPCR comprising a transmembrane 6 region and an intracellular loop 3 region, the

algorithmic approach comprising the steps of:

(a) selecting an endogenous human GPCR comprising a proline residue in the transmembrane-6 region;

(b) identifying, by counting 16 amino acid residues from the proline residue of step (a), in a carboxy-terminus to amino-terminus direction, an endogenous amino acid residue;

(c) altering the identified amino acid residue of step (b) to a non-endogenous amino acid residue to create a non-endogenous version of the endogenous human GPCR; and

(d) determining if the non-endogenous version of the endogenous human GPCR of step (c) is constitutively active.

29. The algorithmic approach of claim 28 wherein the amino acid residue that is two residues from said proline residue in the transmembrane 6 region, in a carboxy-terminus to amino-terminus direction, is tryptophan.

30. A constitutively active, non-endogenous human GPCR produced by the algorithmic approach of claim 28.

31. A constitutively active, non-endogenous human GPCR produced by the algorithmic approach of claim 29.

32. A method for directly identifying a compound selected from the group consisting of inverse agonists, agonists and partial agonists to a non-endogenous, constitutively activated human G protein coupled receptor, said receptor comprising a transmembrane-6 region and an intracellular loop-3 region, comprising the steps of:

(a) selecting an endogenous human GPCR;

(b) identifying a proline residue within the transmembrane-6 region of the GPCR of step (a);

(c) identifying, in a carboxy-terminus to amino-terminus direction, the endogenous, $16^{th}$ amino acid residue from the proline residue of step (b);

(d) altering the endogenous amino acid of step (c) to a non-endogenous amino acid;

(e) confirming that the non-endogenous GPCR of step (d) is constitutively active;

(f) contacting a candidate compound with the non-endogenous, constitutively-activated GPCR of step (e); and

(g) determining, by measurement of the compound efficacy at said contacted receptor, whether said compound is an inverse agonist, agonist or partial agonist of said receptor.

33. The method of claim 32 wherein the non-endogenous amino acid of step (d) is lysine.

34. A compound directly identified by the method of claim 32.

35. The method of claim 32 wherein the directly identified compound is an inverse agonist.

36. The method of claim 32 wherein the directly identified compound is an agonist.--

37. The method of claim 32 wherein the directly identified compound is a partial agonist.

38. A composition comprising the inverse agonist of claim 35.

39. A composition comprising the agonist of claim 36.

40. A composition comprising the partial agonist of claim 37.

41. A method for directly identifying an inverse agonist to a non-endogenous,

constitutively activated human G protein coupled receptor ("GPCR"), said GPCR comprising a transmembrane-6 region and an intracellular loop-3 region, comprising the steps of:

(a) selecting an endogenous human GPCR;

(b) identifying a proline residue within the transmembrane-6 region of the GPCR of step (a);

(c) identifying, in a carboxy-terminus to amino-terminus direction, the endogenous, $16^{th}$ amino acid residue from the proline residue of step (b);

(d) altering the endogenous amino acid of step (c) to a non-endogenous lysine residue;

(e) confirming that the non-endogenous GPCR of step (d) is constitutively active;

(f) contacting a candidate compound with the non-endogenous, constitutively-activated GPCR of step (e); and

(g) determining, by measurement of the compound efficacy at said contacted receptor, whether said compound is an inverse agonist of said receptor.

42. An inverse agonist directly identified by the method of claim 37.

43. A composition comprising an inverse agonist of claim 38.

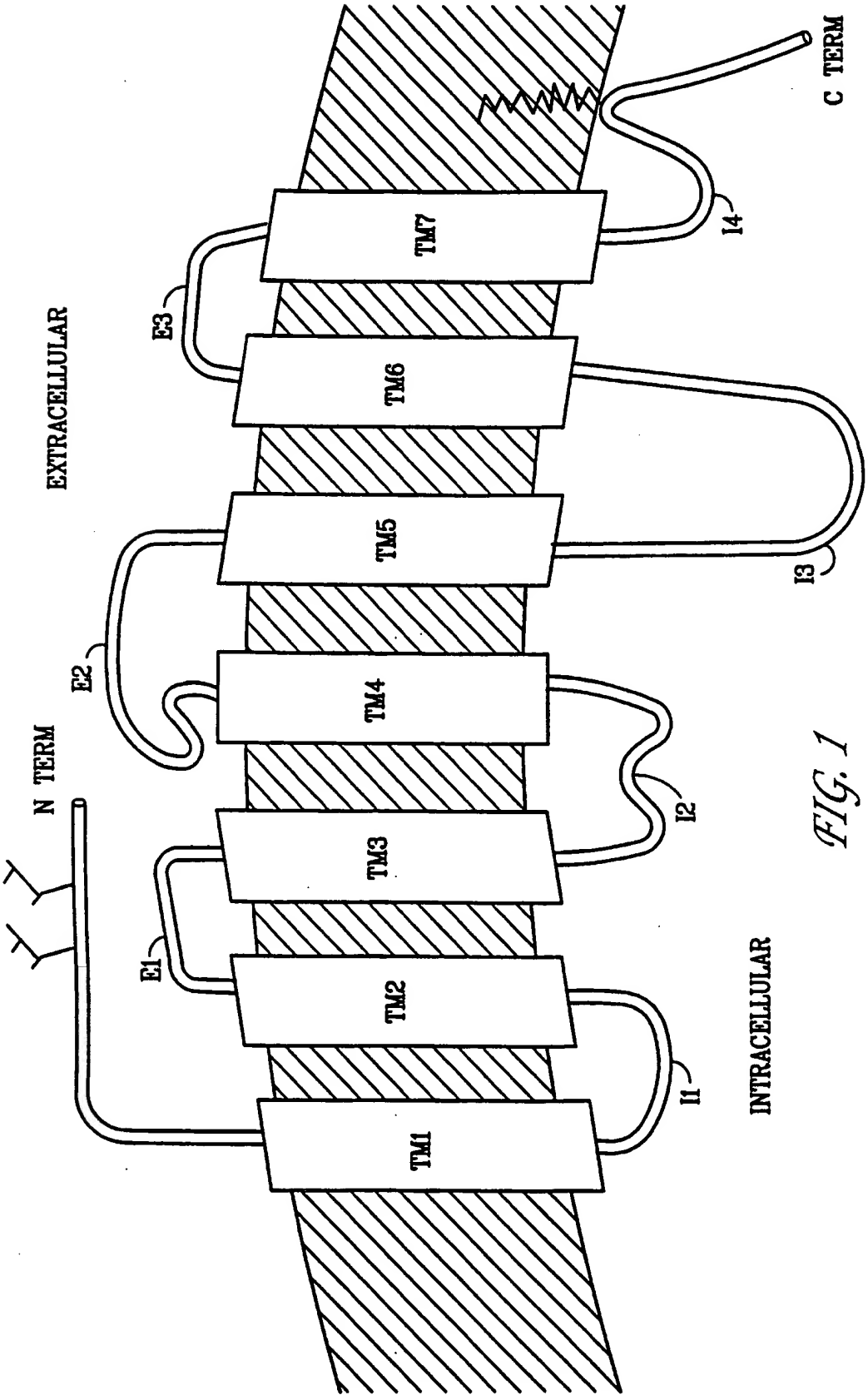

N TERM
C TERM
EXTRACELLULAR
INTRACELLULAR
TM1
TM2
TM3
TM4
TM5
TM6
TM7
E1
E2
E3
I1
I2
I3
I4

FIG. 1

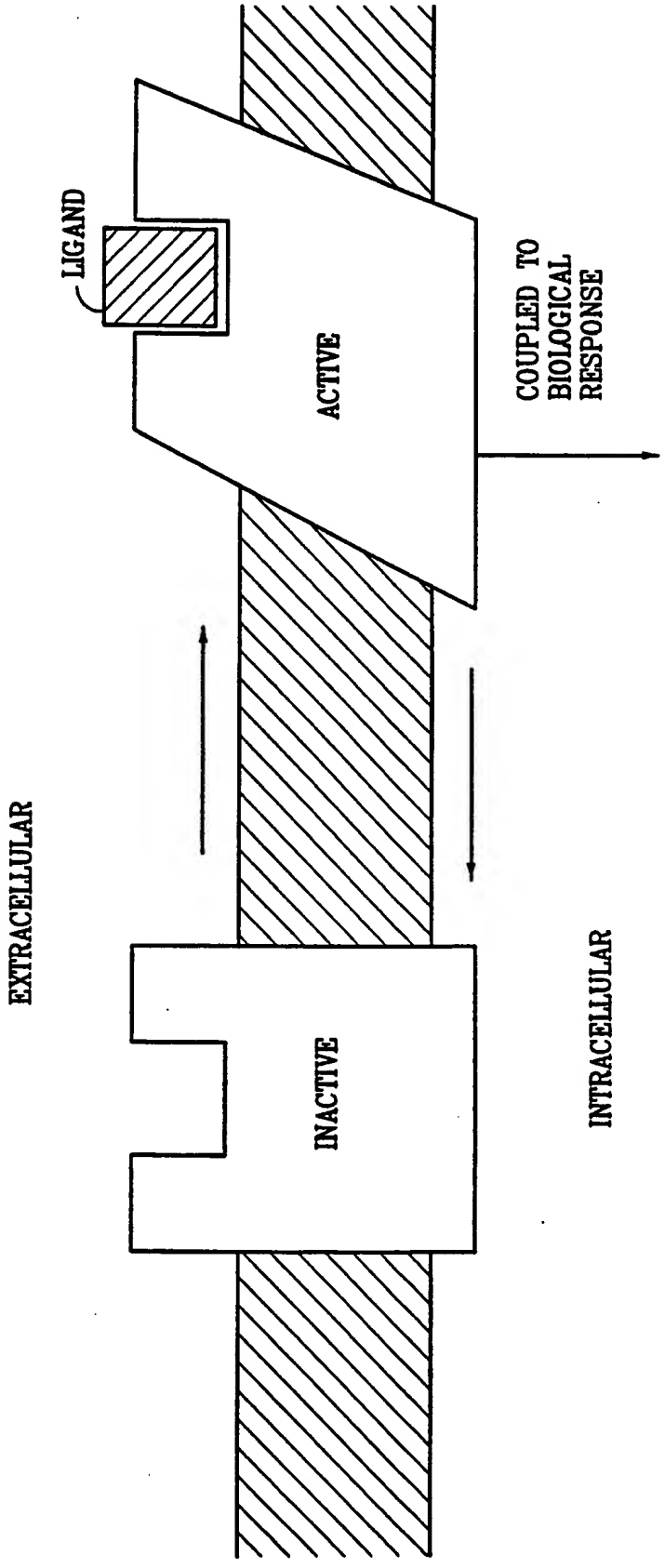

LIGAND
ACTIVE
COUPLED TO BIOLOGICAL RESPONSE
EXTRACELLULAR
INACTIVE
INTRACELLULAR

FIG. 2

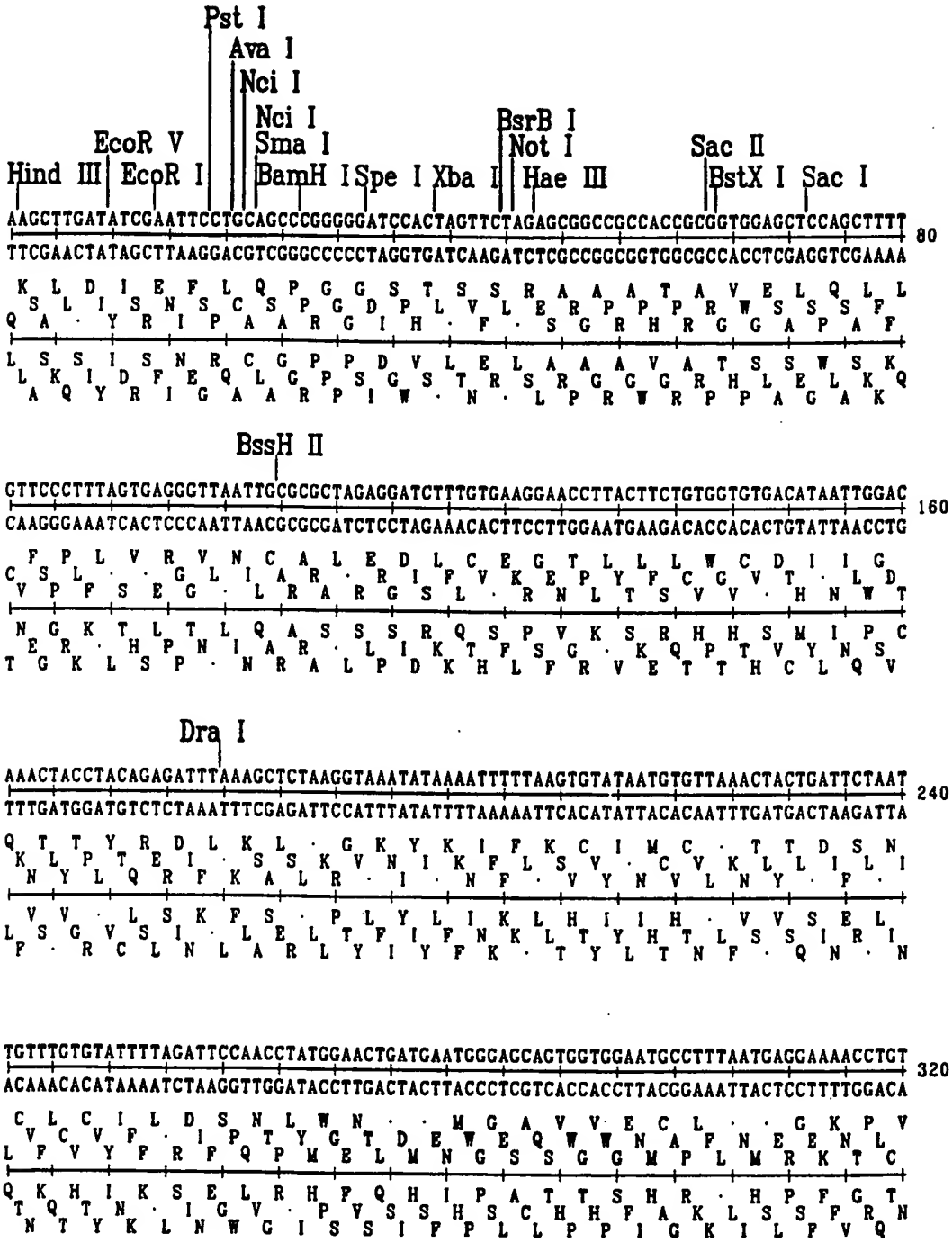

Pst I
Ava I
Nci I
Nci I
Sma I
BamH I
Spe I
Xba I
BsrB I
Not I
Hae III
Sac II
BstX I
Sac I
EcoR V
EcoR I
Hind III
AAGCTTGATATCGAATTCCTGCAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTT
TTCGAACTATAGCTTAAGGACGTCGGGCCCCCTAGGTGATCAAGATCTCGCCGGCGGTGGCGCCACCTCGAGGTCGAAAA
80
BssH II
GTTCCCTTTAGTGAGGGTTAATTGCGCGCTAGAGGATCTTTGTGAAGGAACCTTACTTCTGTGGTGTGACATAATTGGAC
CAAGGGAAATCACTCCCAATTAACGCGCGATCTCCTAGAAACACTTCCTTGGAATGAAGACACCACACTGTATTAACCTG
160
Dra I
AAACTACCTACAGAGATTTAAAGCTCTAAGGTAAATATAAAATTTTTAAGTGTATAATGTGTTAAACTACTGATTCTAAT
TTTGATGGATGTCTCTAAATTTCGAGATTCCATTTATATTTTAAAAATTCACATATTACACAATTTGATGACTAAGATTA
240
TGTTTGTGTATTTTAGATTCCAACCTATGGAACTGATGAATGGGAGCAGTGGTGGAATGCCTTTAATGAGGAAAACCTGT
ACAAACACATAAAATCTAAGGTTGGATACCTTGACTACTTACCCTCGTCACCACCTTACGGAAATTACTCCTTTTGGACA
320

FIG. 3A

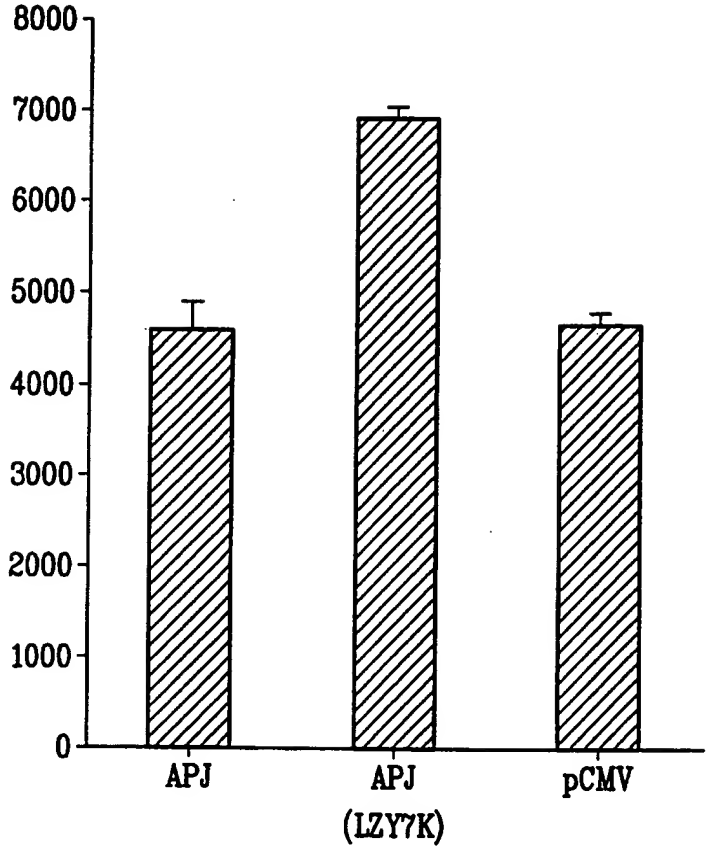
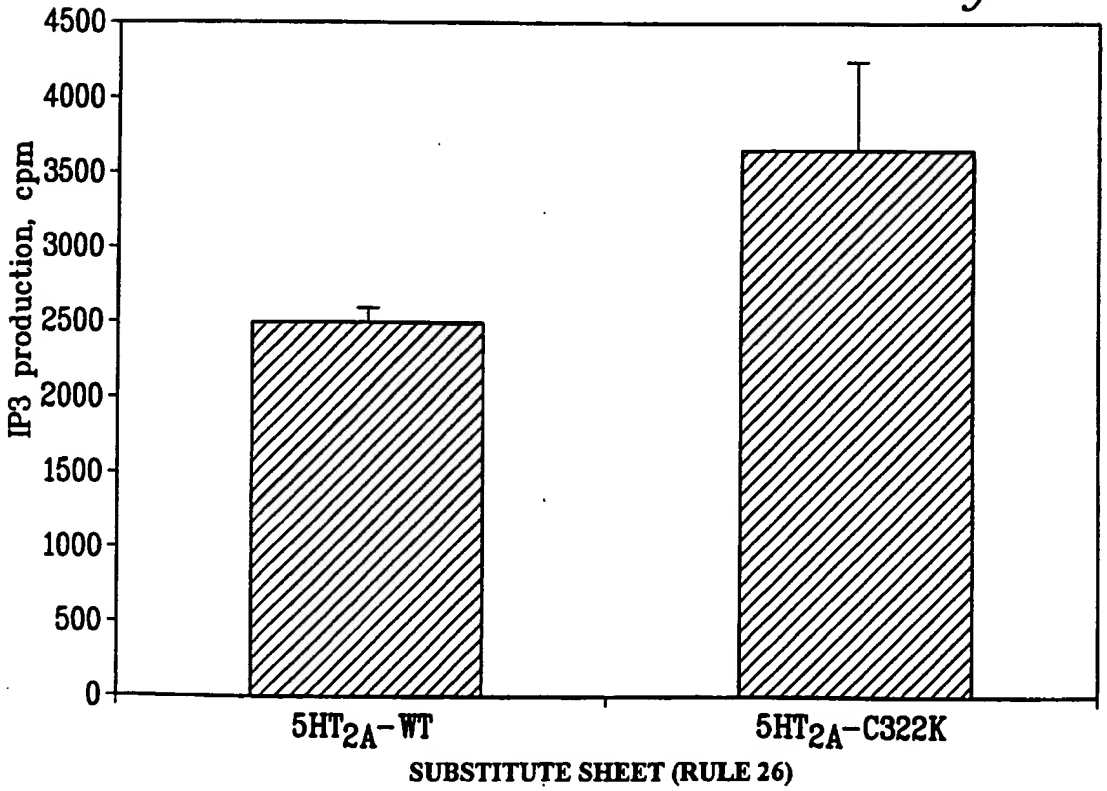

8000
7000
6000
5000
4000
3000
2000
1000
0
APJ
APJ
(LZY7K)
pCMV

FIG. 6

FIG. 7


IP3 production, cpm
4500
4000
3500
3000
2500
2000
1500
1000
500
0
5HT2A-WT
5HT2A-C322K

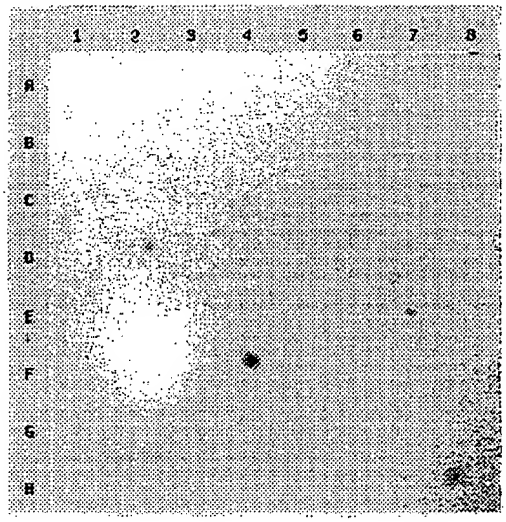



FIG. 8A

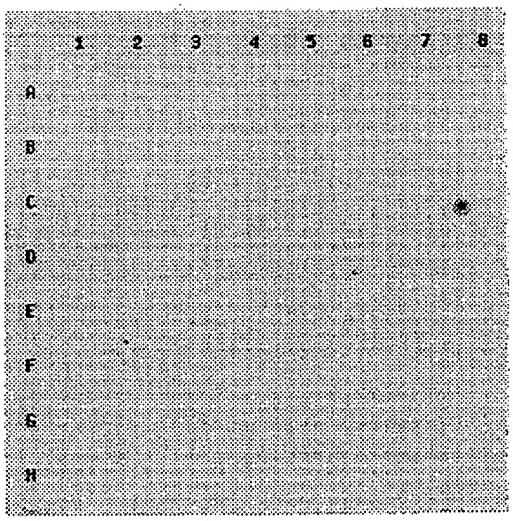



FIG. 8B

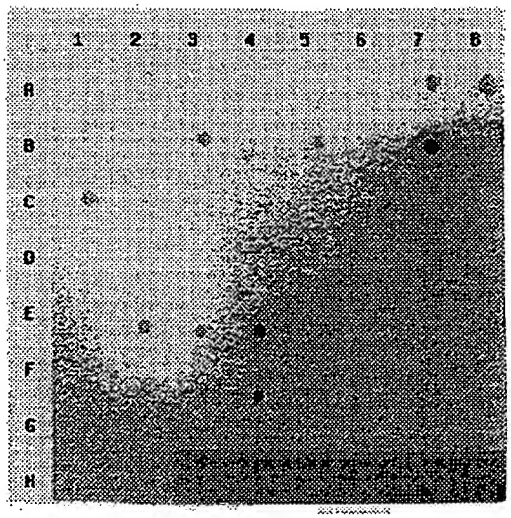



FIG. 8C

# INTERNATIONAL SEARCH REPORT

Int. ional Application No
PCT/US 99/23938

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/72 G01N33/50 G01N33/566

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category ° | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | KJELSBERG M. A. ET AL.: "CONSTITUTIVE ACTIVATION OF THE ALPHA1B-ADRENERGIC RECEPTOR BY ALL AMINO ACID SUBSTITUTIONS AT A SINGLE SITE" JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 267, no. 3, 25 January 1992 (1992-01-25), pages 1430-1433, XP002911764 ISSN: 0021-9258 the whole document --- -/-- | 1,2, 4-13, 15-33, 35-37,41 |

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 2 March 2000 | 09/03/2000 |

| Name and mailing address of the ISA | Authorized officer |
|---|---|
| European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016 | Mandl, B |

Form PCT/ISA/210 (second sheet) (July 1992)

# INTERNATIONAL SEARCH REPORT

Int. ional Application No
PCT/US 99/23938

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication,where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | SCHEER A. ET AL.: "CONSTITUTIVELY ACTIVE G PROTEIN-COUPLED RECEPTORS: POTENTIAL MECHANISMS OF RECEPTOR ACTIVATION" JOURNAL OF RECEPTOR AND SIGNAL TRANSDUCTION RESEARCH, vol. 17, no. 1/03, 1997, pages 57-73, XP000867531 ISSN: 1079-9893 the whole document | 1,2, 4-13, 15-33, 35-37,41 |
| X | WO 97 21731 A (NEW ENGLAND MEDICAL CENTER INC) 19 June 1997 (1997-06-19) the whole document, especially Fig. 2-3 | 1,2,4, 9-13, 20-32, 35-37,41 |
| X | WO 98 38217 A (HERRICK DAVIS KATHARINE ;TEITLER MILT (US); EGAN CHRISTINA C (US)) 3 September 1998 (1998-09-03) the whole document, especially page 7, lines 24-27, and figure 4 | 1,2, 4-13, 15-33, 35-37,41 |
| P,X | PAUWELS P. J. ET AL.: "REVIEW: AMINO ACID DOMAINS INVOLVED IN CONSTITUTIVE ACTIVATION OF G-PROTEIN-COUPLED RECEPTORS" MOLECULAR NEUROBIOLOGY, vol. 17, no. 1/03, 1998, pages 109-135, XP000866477 ISSN: 0893-7648 the whole document | 1,2, 4-13, 15-33, 35-37,41 |

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 99/23938

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 34,38-40,42,43
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
See FURTHER INFORMATION sheet PCT/ISA/210

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

# INTERNATIONAL SEARCH REPORT

International Application No. PCT/US 99 /23938

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 34,38-40,42,43

Claims 34, 38-40, 42 and 43 refer to compounds with an agonistic effect on a GPCR without giving a true technical characterization. Moreover, no such specific compounds are defined in the application. In consequence, the scope of said claims is ambiguous and vague, and their subject-matter is not sufficiently disclosed and supported (Art. 5 and 6 PCT). No search can be carried out for such purely speculative claims whose wording is, in fact, a mere recitation of the results to be achieved.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

# INTERNATIONAL SEARCH REPORT

Information on patent family members

Intl. Ional Application No
PCT/US 99/23938

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---|---|---|
| WO 9721731 A | 19-06-1997 | US 5750353 A | 12-05-1998 |
| | | AU 1334397 A | 03-07-1997 |
| | | CA 2239293 A | 19-06-1997 |
| | | EP 0869975 A | 14-10-1998 |
| WO 9838217 A | 03-09-1998 | AU 6343998 A | 18-09-1998 |

# INTERNATIONAL SEARCH REPORT

International Application No. PCT/US 99 23938

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 34,38-40,42,43

Claims 34, 38-40, 42 and 43 refer to compounds with an agonistic effect on a GPCR without giving a true technical characterization. Moreover, no such specific compounds are defined in the application. In consequence, the scope of said claims is ambiguous and vague, and their subject-matter is not sufficiently disclosed and supported (Art. 5 and 6 PCT).
No search can be carried out for such purely speculative claims whose wording is, in fact, a mere recitation of the results to be achieved.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

# INTERNATIONAL SEARCH REPORT

Information on patent family members

Int. Ional Application No
PCT/US 99/23938

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---|---|---|
| WO 9721731 A | 19-06-1997 | US 5750353 A | 12-05-1998 |
| | | AU 1334397 A | 03-07-1997 |
| | | CA 2239293 A | 19-06-1997 |
| | | EP 0869975 A | 14-10-1998 |
| WO 9838217 A | 03-09-1998 | AU 6343998 A | 18-09-1998 |